

FIGURE 1

GGGGCTTCGGCGCCAGCGGCCAGCGCTAGTCGGTCTGGTAAGGATTTACAAAAGGTGCAGGTATG
AGCAGGTCTGAAGACTAACATTTTGTGAAGTTGTAAAACAGAAAACCTGTTAGAAATGTGGTGGT
TTCAGCAAGGCCTCAGTTTCCTTCCTTCAGCCCTTGTAATTTGGACATCTGCTGCTTTCATATTT
TCATACATTACTGCAGTAACTCCACCATATAGACCCGGCTTACCTTATATCAGTGACACTGG
TACAGTAGCTCCAGAAAAATGCTTATTTGGGGCAATGCTAAATATTGCGGCAGTTTTATGCATTG
CTACCATTTATGTTTCGTATAAGCAAGTTCATGCTCTGAGTCTGAAGAGAACGTTATCATCAAA
TTAAACAAGGCTGGCCTTGTACTTGAATACTGAGTTGTTTAGGACTTTCCTATTGTGGCAAACCTT
CCAGAAAACAACCCTTTTGGCTGCACATGTAAGTGGAGCTGTGCTTACCTTTGGTATGGGCTCAT
TATATATGTTTGTTCAGACCATCCTTTCCTACCAAATGCAGCCCAAATCCATGGCAAACAAGTC
TTCTGGATCAGACTGTTGTTGGTTATCTGGTGTGGAGTAAGTGCACCTTAGCATGCTGACTTGCTC
ATCAGTTTTGCACAGTGGCAATTTTGGGACTGATTTAGAACAGAACTCCATTGGAACCCCGAGG
ACAAAGGTTATGTGCTTCACATGATCACTACTGCAGCAGAATGGTCTATGTCATTTTCCTTCTTT
GGTTTTTTCCTGACTTACATTCGTGATTTTCAGAAAATTTCTTTACGGGTGGAAGCCAATTTACA
TGGATTAACCCCTATGACACTGCACCTTGCCCTATTAACAATGAACGAACACGGCTACTTTCCA
GAGATATTTGATGAAAGGATAAAAATATTTCTGTAATGATTATGATTCTCAGGGATTGGGGAAAGG
TTCACAGAAGTTGCTTATTTCTCTGAAATTTCAACCACCTTAATCAAGGCTGACAGTAACACT
GATGAATGCTGATAATCAGGAAACATGAAAGAAGCCATTTGATAGATTATTTCAAAGGATATCAT
CAAGAAGACTATTAAAAACACCTATGCCTATACTTTTTTATCTCAGAAAATAAAGTCAAAAGACT
ATG

FIGURE 2

<subunit 1 of 1, 266 aa, 1 stop

<MW: 29766, pI: 8.39, NX(S/T): 0

MWWFQQGLSFLPSALVIWTSAAAFIFS YITAVTLHHIDPALPYISDTGTVAPEKCLFGAMLNIAAV
LCIATIYVRYKQVHALSPEENVIIKLNKAGLVLGILSCLGLSIVANFQKTTLFAAHVSGAVLTFG
MGSLYMFVQTIILSYQMOPKIHGKQVFWIRLLLVWCGVSALSMLTCSVVLHSGNFGTDLEQKLHW
NPEDKGYVLHMITTAAEWSMSFSFFGFLLTYIRDFQKISLRVEANLHGLTLYDTAPCPINNERTR
LLSRDI

Important features:

Type II transmembrane domain:

amino acids 13-33

Other Transmembrane domains:

amino acids 54-73, 94-113, 160-180, 122-141

N-myristoylation sites.

amino acids 57-63, 95-101, 99-105, 124-130, 183-189

FIGURE 3

CGGACGCGTGGGCGGACGCGTGGGGGAGAGCCGAGTCCCGGCTGCAGCACCTGGGAGAAGGCAGACC
GTGTGAGGGGGCCTGTGGCCCAGCGTGTGGCCTCGGGGAGTGGGAAGTGGAGGCAGGAGCCCTTC
CTTACACTTCGCCATGAGTTTCCTCATCGACTCCAGCATCATGATTACCTCCCAGATACTATTTTTTG
GATTTGGGTGGCTTTCTTTCATGCGCAATTGTTTAAAGACTATGAGATACGTCAGTATGTTGTACAG
GTGATCTTCTCCGTGACGTTTGCATTTTCTTGACCATGTTTGAGCTCATCATCTTTGAAATCTTAGG
AGTATTGAATAGCAGCTCCCGTTATTTTCACTGAAAAATGAACCTGTGTGTAATCTGCTGATCCTGG
TTTTTCATGGTGCCTTTTTACATTGGCTATTTTTATTGTGAGCAATATCCGACTACTGCATAAAACAACGA
CTGCTTTTTTTCGTCTCTTATGGCTGACCTTTATGTATTCTTCTGGAACTAGGAGATCCCTTTCC
CATTCTCAGCCCAAACATGGGATCTTATCCATAGAACAGCTCATCAGCCGGGTGGTGTGATTGGAG
TGACTCTCATGGCTCTTCTTCTGGATTTGGTGTGTCAACTGCCATACACTTACATGCTTACTTTC
CTCAGGAATGTGACTGACACGGATATTCTAGCCCTGGAACGGGACTGCTGCAAACCATGGATATGAT
CATAAGCAAAAAGAAAAGGATGGCAATGGCAGCGAGAACAATGTTCCAGAAGGGGGAAGTGCATAACA
AACCATCAGGTTTCTGGGGAATGATAAAAAGTGTACCACCTCAGCATCAGGAAGTAAAACTTACT
CTTATCAACAGGAAGTGGATGCTTTGGAAGAATTAAGCAGGCAGCTTTTTCTGAAACAGCTGATCT
ATATGCTACCAAGGAGAGAATAGAATACTCCAAAACCTCAAGGGGAAATATTTAATTTTCTTGGTT
ACTTTTTCTCTATTTACTGTGTTTGGAAAATTTTCATGGCTACCATCAATATTGTTTTTGATCGAGTT
GGGAAAACGGATCCTGTCAAGAGGCATTGAGATCACTGTGAATTATCTGGGAATCCAATTTGATGT
GAAGTTTTGGTCCCAACACATTTCTTCATTCTGTTGGAATAATCATCGTCACATCCATCAGAGGAT
TGCTGATCACTCTTACCAAGTTCTTTTATGCCATCTCTAGCAGTAAGTCTCCAATGTCATTGTCCGT
CTATTAGCACAGATAATGGGCATGTACTTTGTCTCCTCTGTGCTGCTGATCCGAATGAGTATGCCTTT
AGAATACCGCACCATAATCACTGAAGTCCTTGGAGAAGTGCAGTTCAACTTCTATCACCCTGGTTG
ATGTGATCTTCCGGTCAGCGCTCTCTTAGCATACTCTTCTCTATTTGGCTCACAAACAGGCACCA
GAGAAGCAAATGGCACCTTGAACTTAAGCCTACTACAGACTGTTAGAGGCCAGTGGTTTCAAATTTA
GATATAAGAGGGGGGAAAAATGGAACAGGGCCTGACATTTATAAACAACAAAATGCTATGGTAGC
ATTTTTACCTTCATAGCATACTCTTCCCCGTGAGGTGATACTATGACCATGAGTAGCATCAGCCAG
AACATGAGAGGGGAGAACTAAGACAATACTCAGCAGAGAGCATCCCGTGTGGATATGAGGCTGG
TGTAGAGCGGAGAGGAGCAAGAACTAAAGGTGAAAAATACACTGGAACCTGGGGCAAGACATGT
CTATGGTAGCTGAGCCAAACAGTAGGATTTCCGTTTTAAGGTTACATGGAAAAGGTTATAGCTTTG
CCTTGAGATTGACTCATTAAAATCAGAGACTGTAACAAAAAAAAAAAAAAAAAAAAAGGGCGGCCGCG
ACTCTAGAGTCGACCTGCAGAAGCTTGGCCGCCATGGCCCAACTGTTTATTGCAGCTTATAATG

FIGURE 4

MSFLIDSSIMITSQILFFGFGWLFFMRQLFKDYEIRQYVVQVIFSVTFAFSCTMFELIIFEILGV
LNSSSRYPFHWMNLCVILLILVFMVFPFYIGYFIVSNIRLLHKQRLLFSCLLWLFMYFFWKLGD
FPILSPKHGILSIEQLISRVGIVIGVTLMALLSGFGAVNCPYTYMSYFLRNVDTDILALERRLLQ
TMDMIISKKKRMAMARRTMFQKGEVHNKPSGFWGMIKSVTTSASGSENLTLLIQEVDALAEELSRQ
LFLETADLYATKERIEYSKTFKGYFNFLGYFFSIYCVWKIFMATINIVFDRVGKTDVPVTRGIEI
TVNYLGIQFDVKFWSQHISFILVGIIIVTSIRGLLITLTKFFYAISSSKSSNVIVLLLAQIMGM
YFVSSVLLIRMSMPLEYRTIITEVLGELQFNFYHRWFDVIFLVSALSSILFLYLAHKQAPEKQMAP

Important features:

Signal peptide:

amino acids 1-23

Potential transmembrane domains:

amino acids 37-55, 81-102, 150-168, 288-311, 338-356, 375-398,
425-444

N-glycosylation sites.

amino acids 67-70, 180-183 and 243-246

Eukaryotic cobalamin-binding proteins

amino acids 151-160

FIGURE 6

MARCFSLVLLLSIWTRLLVQGLRAEELSIQVSCRIMGITLVSKKANQQLNFTAKEACRLLG
LSLAGKDQVETALKASFETCSYGWVGDFVVISRISPNPKCGKNGVGLIWKVPVSRQFAAYCYN
SSDTWTNSCIPELIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPAPTTTPAPASTSIPR
RKKLICVTEVFMETSTMSTETEPFVENKAAFKNAAAGFGGVPTALLVLLALLFFGAAAGLGFCYVK
RYVKAFPFTNKNQOKEMIETKVVKEEKANDSNPNEESKKTDKNPEESKSPSKTTVRCLEAEV

Signal sequence:

amino acids 1-16

Transmembrane domain:

amino acids 235-254

N-glycosylation site.

amino acids 53-57, 130-134, 289-293

Casein kinase II phosphorylation site.

amino acids 145-149, 214-218

Tyrosine kinase phosphorylation site.

amino acids 79-88

N-myristoylation site.

amino acids 23-29, 65-71, 234-240, 235-239, 249-255, 253-259

FIGURE 7

CGCCGCGCTCCCGCACCCGCGGCCCGCCACCAGCGCGCTCCCGCATCTGCACCCGAGCCCGG
GGCTCCCGGGCGGAGCGAGCAGATCCAGTCCGGCCCGAGCGCAACTCGGTCCAGTCCGGGGCG
CGGCTGCGGGCGCAGAGCGGAGATGCAGCGGCTTGGGGCCACCCTGCTGTGCCTGCTGTGGCG
CGGGGTCCCCACGGCCCCGCGCCGCTCCGACGGCGACCTCGGCTCCAGTCAAGCCCGGCCG
GCTCTCAGTACCCGAGGAGGAGGCCACCCTCAATGAGATGTTCCGCGAGGTTGAGGAAGTATG
GGAGGACACGCAGCACAATTCGCGCAGCGCGGTGGAAGAGATGGAGGCAGAAGAAGCTGCTGCTA
AAGCATCATCAGAAGTGAACCTGGCAAACCTTACCTCCCAGCTATCACAATGAGACCAACACAGAC
ACGAAGGTTGGAAATAATACCATCCATGTGCACCGAGAATTCACAAGATAACCAACAACCAGAC
TGGACAAATGGTCTTTTCAGAGACAGTTATCACATCTGTGGGAGACGAAGAAGGCAGAAGGAGCC
ACGAGTGCATCATCGACGAGGACTGTGGGCCAGCATGTACTGCCAGTTTGCCAGCTTCCAGTAC
ACCTGCCAGCCATGCCGGGGCCAGAGGATGCTCTGCACCCGGGACAGTGAAGTGTGGAGACCA
GCTGTGTGTCTGGGGTCACTGCACCAAAATGGCCACCAGGGGCAGCAATGGGACCATCTGTGACA
ACCAGAGGGACTGCCAGCCGGGGCTGTGCTGTGCCCTTCCAGAGAGGCTGCTGTTCCCTGTGTGC
ACACCCCTGCCCGTGGAGGGCGAGCTTTGCCATGACCCCGCAGCCGGCTTCTGGACCTCATCAC
CTGGGAAATAGAGCTGATGGAGCCTTGACCGATGCCCTTGTGCCAGTGGCCCTCCTCTGCCAGC
CCCACAGCCACAGCCTGGTGTATGTGTGCAAGCCGACCTTCGTGGGGAGCCGTGACCAAGATGGG
GAGATCCTGCTGCCAGAGAGGTCCCCGATGAGTATGAAGTTGGCAGCTTCATGGAGGAGGTGCG
CCAGGAGCTGGAGGACCTGGAGAGGAGCCTGACTGAAGAGATGGCGTGGGGGAGCCTGCGGCTG
CCGCCGCTGCACTGCTGGGAGGGGAAGAGATTAGATCTGGACCAGGCTGTGGGTAGATGTGCAA
TAGAAATAGCTAATTTATTTCCCCAGGTGTGTGCTTTAGGCGTGGGCTGACCAGGCTTCTTCTA
CATCTTCTTCCAGTAAGTTTCCCCTCTGGCTTGACAGCATGAGGTGTTGTGCATTTGTTAGCT
CCCCCAGGCTGTTCTCCAGGCTTCCAGTCTGGTGTGTTGGGAGAGTCAGGCAGGTTAAACTGCA
GGAGCAGTTTCCACCCCTGTCCAGATTATTGGCTGCTTTGCCCTTACCAGTTGGCAGACAGCCG
TTGTTCTACATGGCTTTGATAATTGTTTGGAGGGAGGAGATGGAAACAATGTGGAGTCTCCCTC
TGATTGGTTTTGGGAAATGTGGAGAAGAGTGCCCTGCTTTGCAAACATCAACCTGGCAAAAATG
CAACAAATGAATTTCCACGCAGTCTTTCCATGGGCATAGGTAAGCTGTGCCCTTCCAGCTGTGTC
AGATGAAATGTTCTGTTACCCCTGCATTACATGTGTTTATTTCATCCAGCAGTGTGCTCAGCTCC
TACCTCTGTGCCAGGGCAGCATTTCATATCCAAGATCAATTCCTCTCTCAGCACAGCCTGGGG
AGGGGGTCATTGTTCTCCTCGTCCATCAGGGATCTCAGAGGCTCAGAGACTGCAAGCTGCTTGCC
CAAGTCACACAGCTAGTGAAGACCAGAGCAGTTTCATCTGGTTGTGACTCTAAGCTCAGTGTCT
CTCCACTACCCACACCAGCCTTGGTGCCACAAAAGTGTCTCCCAAAAGGAAGGAGAATGGGAT
TTTTCTTGAGGCATGCACATCTGGAATTAAGGTCAAACCTAATCTCACATCCCTCTAAAAGTAAA
CTACTGTTAGGAACAGCAGTGTCTCACAGTGTGGGGCAGCCGTCCTTCTAATGAAGACAATGAT
ATTGACACTGTCCCTCTTTGGCAGTTGCATTAGTAACTTTGAAAGGTATATGACTGAGCGTAGCA
TACAGGTTAACCTGCAGAAACAGTACTTAGGTAATTGTAGGGCGAGGATTATAAATGAAATTTGC
AAAATCACTTAGCAGCAACTGAAGACAATTATCAACCACGTGGAGAAAATCAAACCGAGCAGGGC
TGTGTGAAACATGGTTGTAATATGCGACTGCGAACACTGAACTCTACGCCACTCCACAAATGATG
TTTTCAGGTGTCATGGACTGTTGCCACCATGTATTCATCCAGAGTCTTAAAGTTTAAAGTTGCA
CATGATTGTATAAGCATGCTTTCTTTGAGTTTTAAATTATGTATAAACATAAGTTGCATTTAGAA
ATCAAGCATAAATCACTTCAACTGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 8

MQRLGATLLCLLLAAAVPTAPAPAPTATSAPVKPGPALSYPQEEATLNEMFREVEELMEDTQHKL
RSAVEEMEAEEAAKASSEVNLANLPPSYHNETNTDTKVGNNTIHVHREIHKITNNQTGMVFSE
TVITSVGDEEGRRSHECIIDEDCGPSMYCQFASFYTCQPCRQRMCLTRDSECCGDQLCVWGH
TKMATRGSNGTICDNQRDCQPLCCAFQRGLLFPVCTPLPVEGELCHDPASRLLDLITWELEPDG
ALDRPCASGLLCQPHSHSLVYVCKPTFVGSRDQDGEILLPREVPDEYEVGSFMEEVRQELEDLE
RSLTEEMALGEPAAAAAALLGGEEI

Signal sequence:

amino acids 1-19

N-glycosylation site.

amino acids 96-100, 106-110, 121-125, 204-208

Casein kinase II phosphorylation site.

amino acids 46-50, 67-71, 98-102, 135-139, 206-210, 312-316,
327-331

N-myristoylation site.

amino acids 202-208, 217-223

Amidation site.

amino acids 140-144

FIGURE 9

CGGACGCGTGGGCGGACGCGTGGGGCTGTGAGAAAGTGCCAATAAATACATCATGCAACCCAC
GGCCACCTTGTGAACCTCGTGCCAGGGCTGATGTGCGTCTTCCAGGGCTACTCATCCAAAG
GCCTAATCCAACGTTCTGTCTTCAATCTGCAAATCTATGGGGTCTGGGGCTTCTTGACCCTT
AACTGGGTACTGGCCCTGGGCAATGCGTCTCGCTGGAGCCTTTGCCTCCTTCTACTGGGCCTT
CCACAAGCCCCAGGACATCCCTACCTTCCCCTTAATCTCTGCCTTCATCCGCACACTCCGTTACC
ACACTGGGTCAATTGGCATTGGAGCCCTCATCCTGACCCTTGTGCAGATAGCCCGGGTCATCTG
GAGTATATTGACCACAAGCTCAGAGGAGTGCAGAACCCTGTAGCCCGCTGCATCATGTGCTGTTT
CAAGTGCTGCCTCTGGTGTCTGAAAAATTTATCAAGTTCCTAAACCGCAATGCATACATCATGA
TCGCCATCTACGGGAAGAATTTCTGTGTCTCAGCCAAAAATGCGTTCATGCTACTCATGCGAAAC
ATTGTCAGGGTGGTCTGCTCTGGACAAAGTACAGACCTGCTGCTGTTCTTTGGGAAGCTGCTGGT
GGTCGGAGGCGTGGGGTCTGTCTTCTTTTTTCTCCGGTGCATCCCGGGGCTGGGTAAAG
ACTTTAAGAGCCCCACCTCAACTATTACTGGCTGCCATCATGACCTCCATCCTGGGGCCTAT
GTCATCGCCAGCGGCTTCTTCAGCGTTTTTCGGCATGTGTGTGGACACGCTCTTCCTCTGCTTCCT
GGAAGACCTGGAGCGGAACAACGGCTCCCTGGACCGGCCCTACTACATGTCCAAGAGCCTTCTAA
AGATTCTGGGCAAGAAGAACGAGGCGCCCCGGACAACAAGAAGAGGAAGAAGTGACAGCTCCGG
CCCTGATCCAGGACTGCACCCACCCACCGTCCAGCCATCCAACCTCACTTCGCCTTACAGGT
CTCCATTTTGTGGTAAAAAAGTTTTAGGCCAGGCGCCGTGGCTCACGCCTGTAATCCAACACT
TTGAGAGGCTGAGGCGGGCGGATCACCTGAGTCAGGAGTTCGAGACCAGCCTGGCCAACATGGTG
AAACCTCCGTCTCTATTAATAAATACAAAAATTAGCCGAGAGTGGTGGCATGCACCTGTCATCCCA
GCTACTCGGGAGGCTGAGGCAGGAGAAATCGCTTGAACCCGGGAGGCAGAGGTTGCAGTGAGCCGA
GATCGCGCCACTGCACCTCAACCTGGGTGACAGACTCTGTCTCCAAAACAAAACAAACAA
AAGATTTTATTAAAGATATTTGTTAACCTC

FIGURE 10

RTRGRTRGGCEKVPINTSCNPTAHLVNSSCPGLMCFVQGYSSKGLIQRSVFNLQIYGVLGLEFWTL
NWWLALGQCVLGAFASFYWAFHKPQDIPTFPLISAFIRTLRYHTGSLAFGALILTLVQIARVIL
EYIDHKLRGVQNPVARCIMCCFKCCLWCLEKFIKFLNRNAYIMIAIYGKNFCVSAKNAFMLLMRN
IVRVVLDKVTDLLLFFGKLLVVGGVVLSFFFFSGRIPGLGKDFKSPHLNYYWLPIMTSLGAY
VIASGFFSVFGMCDTLFLCFLEDLERNNGSLDRPYMSKSLKILGKKNEAPPDNKKRKK

Important features:

Transmembrane domains:

amino acids 57-80 (type II), 110-126, 215-231, 254-274

N-glycosylation sites.

amino acids 16-20, 27-31, 289-293

Hypothetical YBR002c family proteins.

amino acids 276-288

Ammonium transporters proteins.

amino acids 204-231

N-myristoylation sites.

amino acids 60-66, 78-84

Amidation site.

amino acids 306-310

FIGURE 12

MGACLGACSLSCASCLCGSAPCILCSCCPASRNSTVSRLI FTFFLFLGVLVSI IMLSPGVESQL
YKLPWVCEEAGIPTVLQGHIDCGSLLGYRAVYRMCFATAAFFFFFFFFLLMLCVSSSRDPRAAIQ
NGFWFFKFLILVGLTVGAFYIPDGSFTNIWFYFGVVGSFLFILIQLVLLIDFAHSWNQRWLGKAE
ECDSRAWYAGLFFFFFFLLFYLLSIAAVALMFMYYTEPSGCHEGKVFISLNLTFVCVSIAAVLPKV
QDAQPNSGLLQASVITLYTMFVTWSALSSIPEQKCNPHLPTQLGNETVVAGPEGYETQWWDAPSI
VGLIIFLLCTLFISLRSSDHRQVNSLMQTEECPPMLDATQQQQQQVAACEGRAFDNEQDGVITYSY
SFFHFCLV LASLHVMMTLTNWYKPGETRKMISTWTAVVVKICASWAGLLLYLWTLVAPLLLRNRD
FS

Signal sequence:

amino acids 1-20

Transmembrane domains:

amino acids 40-58, 101-116, 134-150, 162-178, 206-223, 240-257,
272-283, 324-340, 391-406, 428-444

FIGURE 14

MNHLPEDMENALTGSQSSHASLRNIHSINPTQLMARIESYEGREKKGISDVRRTFCLFVTFDLLF
VTLLWIIELNVNGGIENTLEKEVMQYDYSSYFDIFLLAVFRFKVLILAYAVCRLRHHWAIALTT
AVTSAFLLAKVILSKLFSQGAFGYVLPPIISFILAWIETWFLDFKVLQPQEAEEENRLLIVQDASER
AALIPGGLSDGQFYSPPESEAGSEEAEEKQDSEKPLEL

Important features of the protein:

Signal peptide:

amino acids 1-20

Transmembrane domains:

amino acids 54-72, 100-118, 130-144, 146-166

N-myristoylation sites.

amino acids 14-20, 78-84, 79-85, 202-208, 217-223

FIGURE 15

ACTCGAACGCAGTTGCTTCGGGACCCAGGACCCCTCGGGCCCGACCCGCCAGGAAAGACTGAGG
CCGCGGCCTGCCCGCCCGGCTCCCTGCGCCGCCCGCCTCCCGGGACAGAAGATGTGCTCCAG
GGTCCCTCTGCTGCTGCCGTGCTCCTGCTACTGCCCTGGGGCTGGGGTGCAGGGCTGCCCAT
CCGGCTGCCAGTGCAGCCAGCCACAGACAGTCTTCTGCACTGCCCGCCAGGGGACCACGGTGCCC
CGAGACGTGCCACCCGACACGGTGGGGCTGTACGTCTTTGAGAACGGCATCACCATGCTCGACGC
AGGCAGCTTTGCCGGCCTGCCGGCCCTGCAGCTCCTGGACCTGTACAGAACCAGATCGCCAGCC
TGCCAGCGGGGTCTCCAGCCACTCGCCAACCTCAGCAACCTGGACCTGACGGCCAACAGGCTG
CATGAAATCACCAATGAGACCTTCCGTGGCCTGCGGCGCCTCGAGCGCCTCTACCTGGGCAAGAA
CCGCATCCGCCACATCCAGCCTGGTGCCTTCGACACGCTCGACCGCCTCCTGGAGCTCAAGCTGC
AGGACAACGAGCTGCGGGCACTGCCCGCCTGCGCCTGCCCGCCTGCTGCTGCTGGACCTCAGC
CACAACAGCCTCCTGGCCTGGAGCCCGGCATCCTGGACACTGCCAACCTGGAGGCGCTGCGGCT
GGCTGGTCTGGGGCTGCAGCAGCTGGACGAGGGGCTCTTACGCCGCTTGCGCAACCTCCACGACC
TGGATGTGTCCGACAACCAGCTGGAGCGAGTGCCACCTGTGATCCGAGGCTCCGGGGCTGACG
CGCTGCGGCTGGCCGCAACACCCGATTGCCAGCTGCGGCCCGAGGACCTGGCCGGCTGGC
TGCCCTGCAGGAGCTGGATGTGAGCAACCTAAGCCTGCAGGCCCTGCCTGGCGACCTCTCGGGCC
TCTTCCCCCGCCTGCGGCTGCTGGCAGCTGCCCGCAACCCCTTCAACTGCGTGTGCCCCCTGAGC
TGGTTTGGCCCTGGGTGCGGAGAGCCACGTCACTGGCCAGCCCTGAGGAGACGCGCTGCCA
CTTCCCGCCCAAGAACCTGGCCGGCTGCTCCTGGAGCTTGACTACGCCACTTTGGCTGCCAG
CCACCACCACAGCCACAGTGCCACCACGAGGCCCGTGGTGGGGAGCCACAGCCTTGCT
TCTAGCTTGGCTCCTACTGGCTTAGCCCCACAGCGCCGGCCACTGAGGCCCCAGCCCGCCCTC
CACTGCCCCACCAGCTGTAGGGCCTGTCCCCAGCCCCAGGACTGCCACCCTCCACCTGCCCA
ATGGGGGCACATGCCACCTGGGGACACGGCACCCCTGGCGTGCTTGTGCCCCGAAGGCTTACG
GGCCTGTACTGTGAGAGCCAGATGGGGCAGGGGACAGGCCACCCCTACACCAGTACGCCGAG
GCCACCACGGTCCCTGACCCTGGGCATCGAGCCGGTGAAGCCACCTCCCTGCGCGTGGGGCTGC
AGCGCTACCTCCAGGGGAGCTCCGTGCAGCTCAGGAGCCTCCGTCTACCTATCGCAACCTATCG
GGCCCTGATAAGCGGCTGGTGAAGCTGCGACTGCCTGCCTCGCTCGCTGAGTACAGGTCACCCA
GCTGCGGCCCAACGCCACTTACTCCGTCTGTGTATGCCTTTGGGGCCCGGGGGTGCCTGAGG
GCGAGGAGGCTGCGGGGAGGCCATACACCCAGCCGCTCCACTCCAACCACGCCCCAGTCACC
CAGGCCCGGAGGGCAACCTGCCGCTCCTCATTGCGCCCGCCTGGCCGCGGTGCTCCTGGCCGC
GCTGGCTGCGGTGGGGGACGCTACTGTGTGCGGCGGGGGCCATGGCAGCAGCGGCTCAGG
ACAAAGGCGAGGTGGGGCCAGGGCTGGGCCCTGGAACCTGGAGGGAGTGAAGGTCCCTTGGAG
CCAGGCCCAAGGCAACAGAGGGCGGTGGAGAGGCCCTGCCACGCGGCTGAGTGTGAGGTGCC
ACTCATGGGCTTCCCAGGGCCTGGCCTCCAGTACCCCTCCACGCAAAGCCCTACATCTAAGCCA
GAGAGAGACAGGGCAGCTGGGGCCGGCTCTCAGCCAGTGAAGTGGCCAGCCCCCTCCTGTGCC
ACACCACGTAAGTTCTCAGTCCCAACCTCGGGGATGTGTGCAGACAGGGCTGTGTGACCACAGCT
GGGCCCTGTCCCTCTGGACCTCGGTCTCCTCATCTGTGAGATGCTGTGGCCAGCTGACGAGCC
CTAACGTCCCCAGAACCGAGTGCTATGAGGACAGTGTCCGCCCTGCCCTCCGCAACGTGCAGTC
CCTGGGCACGGCGGGCCCTGCCATGTGCTGGTAAAGCATGCCTGGGTCTGCTGGGCTCTCCAC
TCCAGGCCGACCCTGGGGCCAGTGAAGGAAGCTCCCGAAAGAGCAGAGGGAGAGCGGGTAGGC
GGCTGTGTGACTCTAGTCTTGGCCCCAGGAAGCAAGGAACAAAAGAACTGGAAAGGAAGATGC
TTTAGGAACATGTTTTGCTTTTTTAAAAATATATATATTTATAAGAGATCCTTTCCATTATTCT
GGGAGATGTTTTTCAAACCTCAGAGACAAGGACTTTGGTTTTTGTAAAGCAAACGATGATATGAA
GGCCTTTTGTAAAGAAAAATAAAAGATGAAGTGTAAA

10063728

FIGURE 17

GCAGCGGCGAGGGCGGGTGGTGGCTGAGTCCGTGGTGGCAGAGGCGAAGGCCAGAGCTCATGCG
GGTCCGGATAGGGCTGACGCTGCTGCTGTGTGCGGTGCTGCTGAGCTTGGCCTCGGCGTCTCGG
ATGAAGAAGGCAGCCAGGATGAATCCTTAGATTCCAAGACTACTTTGACATCAGATGAGTCAGTA
AAGGACCATACTACTGCAGGCAGAGTAGTTGCTGGTCAAATATTTCTTGATTAGAAGAATCTGA
ATTAGAATCCTCTATTCAAGAAGAGGAAGACAGCCTCAAGAGCCAAGAGGGGAAAGTGTACAG
AAGATATCAGCTTTCTAGAGTCTCCAAATCCAGAAAACAAGGACTATGAAGAGCCAAAGAAAAGTA
CGGAAACCAGCTTTGACCGCCATTGAAGGCACAGCACATGGGGAGCCCTGCCACTTCCCTTTTCT
TTTCTAGATAAGGAGTATGATGAATGTACATCAGATGGGAGGGAAGATGGCAGACTGTGGTGTG
CTACAACCTATGACTACAAAGCAGATGAAAAGTGGGGCTTTTGTGAAACTGAAGAAGAGGCTGCT
AAGAGACGGCAGATGCAGGAAGCAGAAATGATGTATCAAACCTGGAATGAAAATCCTTAATGGAAG
CAATAAGAAAAGCCAAAAAGAGAAGCATATCGGTATCTCCAAAAGGCAGCAAGCATGAACCATA
CCAAAGCCCTGGAGAGAGTGTACATATGCTCTTTTATTTGGTGATTACTTGCCACAGAATATCCAG
GCAGCGAGAGAGATGTTTGAGAAGCTGACTGAGGAAGGCTCTCCCAAGGGACAGACTGCTCTTGG
CTTTCTGTATGCCTCTGGACTTGGTGTAAATTCAAGTCAGGCAAAGGCTCTTGTATATTATACAT
TTGGAGCTCTTGGGGCAATCTAATAGCCACATGGTTTTGGTAAGTAGACTTTAGTGGAAGGCT
AATAATATTAACATCAGAAGAATTTGTGGTTTATAGCGGCCACAACCTTTTCAGCTTTCATGATC
CAGATTTGCTTGTATTAAGACCAAATATTCAGTTGAACTTCCCTCAAATCTTGTAAATGGATAT
AACACATGGAATCTACATGTAAATGAAAGTTGGTGGAGTCCACAATTTTTCTTTAAATGATTAG
TTTGGCTGATTGCCCCAAAAAGAGAGATCTGATAAATGGCTCTTTTTAAATTTTCTCTGAGTTG
GAATTGTCAGAATCATTTTTTACATTAGATTATCATAATTTTAAAAATTTTTCTTTAGTTTTTCA
AAATTTTGTAAATGGTGGCTATAGAAAAACAACATGAAATATTATACAATATTTTGCAACAATGC
CCTAAGAATTTGTTAAATTCATGGAGTTATTTGTGCAGAATGACTCCAGAGAGCTCTACTTTCTG
TTTTTTACTTTTCATGATTGGCTGTCTTCCATTTATTCTGGTCATTTATTGCTAGTGACACTGT
GCCTGCTCCAGTAGTCTCATTTTCCCTATTTGCTAATTTGTTACTTTTTCTTTGCTAATTTGG
AAGATTAAGTCAATTTTAAATAAATTTATGTCTAAGATTAAAAAAAAAAAAAAAAAAAAAAAAAA
AA

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

FIGURE 18

MRVRIGLTLILLCAVLLSLASASSDEEGSQDESLSKTTLTSDSVKDHTTAGRVVAGQIFLDSESEL
ESSIQEEEDSLKSQEGESVTEDISFLESPNPENKDYEEPKKVRKPALTAIEGTAHGEPCHFPFLFDK
EYDECTSDGREDGRLWCATTYDYKADEKWFCETEEEAAKRRMQEAEMMYQTGMKILNGSNKKSQKR
EAYRYLQKAASMNHTKALERSYALLFGDYLPQNIQAAREMFELTEEGSPKGQTALGFLYASGLGVN
SSQAKALVYYTFGALGGNLI AHMVLVSRL

Important features:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 195-199, 217-221, 272-276

Tyrosine kinase phosphorylation site.

amino acids 220-228

N-myristoylation sites.

amino acids 120-126, 253-259, 268-274, 270-274, 285-291, 289-295

Glycosaminoglycan attachment site.

amino acids 267-271

Microbodies C-terminal targeting signal.

amino acids 299-303

Type II fibronectin collagen-binding domain protein.

amino acids 127-169

Fructose-bisphosphate aldolase class-II protein.

amino acids 101-119

FIGURE 20

MLFWVLGLLILCGFLWTRKGLKIEDITDKYIFITGCDSGFGNLAARTFDKKGPHVIAACLTESG
STALKAETSERLRTVLLDVTDPENVKRTAQWVKNQVGEKGLWGLINNAGVPGVLAPTDWLTLEDY
REPIEVNLFGLISVTLNMLPLVKKAQGRVINVSSVGGRLAIVGGGYTPSKYAVEGFNDSLRRDMK
AFGVHVSCIEPGLFKTNLADPVKVIIEKKLAIWEQLSPDIKQYEGEGYIEKSLDKLKGKNSYVNMD
LSPVVECMDHALTSLFPKTHYAAGKDAKIFWIPLSHMPAALQDFLLLKQKAE LANPKAV

Important features of the protein:

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 136-152

N-glycosylation sites.

amino acids 161-163, 187-190 and 253-256

Glycosaminoglycan attachment site.

amino acids 39-42

N-myristoylation sites.

amino acids 36-41, 42-47, 108-113, 166-171, 198-203 and 207-212

FIGURE 22

MEGESTSAVLSGFVLGALAFQHLNTSDTEGFLLGVEVKGEAKNSITDSQMDDVEVVYTIIDIQYI
PCYQLFSFYNSSGEVNEQALKKILSNVKNVVGWYKFRHSDQIMTFRERLLHKNLQEHFSNODL
VFLLLTPSIITESCSTHRLHSLYKPKGLFHRVPLVVANLGMSEQLGYKTVSGSCMSTGFSSRAV
QTHSSKFFEEDGSLKEVHKINEMYASLQEELKSICKKVEDSEQAVDKLVKDVNRLKREIEKRRGA
QIQAAAREKNIQKDPQENIFLCQALRTFFPNSEFLHSCVMSLKNRHVSKSSCNYNHHLDVVDNLTL
MVEHTDIPEASPASTPQIIKHKALDLLDRWQFKRSRLLDTDQDKRSKANTGSSNQDKASKMSSPET
DEEIEKMKGFGEYSRSPTF

Important features:

Signal peptide:

amino acids 1-19

N-glycosylation sites.

amino acids 75-79, 322-326

N-myristoylation site.

amino acids 184-154

Growth factor and cytokines receptors family.

amino acids 134-150

FIGURE 24

MARFGLPALLCTLAVLSAALLAELKSKSCSEVRRRLYVSKGFNKNNDAPLHEINGDHLKICPQGST
CCSQEMEEKYSLQSKDDFKSVVSEQCNHLQAVFASRYKKFDEFFKELLENAEKSLNDMFVKTYGH
LYMQNSELFKDLFVELKRYVVGNVNLEEMLNDFWARLLERMFRVNSQYHFTDEYLECVSKYTE
QLKPFQDVPRKLLQVTRAFVAARTFAQGLAVAGDVVSKVSVVNPTAQCTHALLKMIYCSHRGL
VTVKPCYNYCSNIMRGCLANQGDLD FEWNNFIDAMLMVAERLEGPFNIESVMDPIDVKISDAIMN
MQDNSVQVSQKVFQCGPPKPLPAGRISRSISESAFSARFRPHHEERPTAAGTSLDRLVTDVK
EKLKQAKKFWSSLPNVCNDRMAAGNEDDCWNGKGSRYLFAVTGNGLANQGNNPEVQVDTS
KPDILILRQIMALRVMTSKMKNAYNGNDVDFDISDESSGEGSGSGCEYQQCPSEFDYNATDHAG
KSANEKADSAGVRPGAQAYLLTVFCILFLVMQREWR

Important features:

Signal peptide:

amino acids 1-22

ATP/GTP-binding site motif A (P-loop).

amino acids 515-524

N-glycosylation site.

amino acids 514-518

Glycosaminoglycan attachment sites.

amino acids 494-498, 498-502

N-myristoylation sites.

amino acids 63-69, 224-230, 276-282, 438-444, 497-503, 531-537

Glypicans proteins.

amino acids 54-75, 105-157, 238-280, 309-346, 423-460, 468-506

FIGURE 25

CTCGCCCTCAAATGGGAACGCTGGCCTGGGACTAAAGCATAGACCACCAGGCTGAGTATCCTGAC
CTGAGTCATCCCCAGGGATCAGGAGCCTCCAGCAGGGAACCTTCCATTATATTCTTCAAGCAACT
TACAGCTGCACCGACAGTTGCGATGGAAAGTTCTAATCTCTTCCCTCCTCTGTTGCTGCCACTAA
TGCTGATGTCCATGGTCTCTAGCAGCCTGAATCCAGGGTCCGACAGGGCCACAGGGACCGAGGC
CAGGCTTCTAGGAGATGGCTCCAGGAAGGCGCCAAGAATGTGAGTGCAAAGATTGGTTCCTGAG
AGCCCCGAGAAGAAAATTCATGACAGTGTCTGGGCTGCCAAGAAGCAGTGCCCCCTGTGATCATT
TCAAGGGCAATGTGAAGAAAACAAGACACCAAGGCACCACAGAAAGCCAACAAGCATTCCAGA
GCCTGCCAGCAATTTCTCAAACAATGTCAGCTAAGAAGCTTTGCTCTGCCTTTGTAGGGAGCTCTG
AGCGCCCACTCTTCCAATTAACATTCTCAGCCAAGAAGACAGTGAGCACACCTACCAGACACTC
TTCTTCTCCCACCTCACTCTCCCACGTGACCCACCCCTAAATCATTCCAGTGCTCTCAAAAAGCA
TGTTTTTCAAGATCATTTTGTGGTGGTCTCTCTAGTGTCTTCTTCTCTCGTCAGTCTTAGCCT
GTGCCCTCCCCTTACCCAGGCTTAGGCTTAATTACCTGAAAGATTCCAGGAACTGTAGCTTCCT
AGCTAGTGTCAATTAACCTTAAATGCAATCAGGAAAGTAGCAAACAGAAGTCAATAAATATTTTT
AAATGTCAAAAAAAAAAAAAAAAAA

FIGURE 26

MKVLISL L L L L L L L L L P L M L M S M V S S L N P G V A R G H R D R G Q A S R R W L Q E G G Q E C E C K D W F L R A P R R K F M
T V S G L P K K Q C P C D H F K G N V K K T R H Q R H H R K P N K H S R A C Q Q F L K Q C Q L R S F A L P L

Important features:

Signal peptide:

amino acids 1-22

N-myristoylation sites.

amino acids 27-33, 46-52

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52

FIGURE 27

GGACGCCAGCGCCTGCAGAGGCTGAGCAGGGAAAAAGCCAGTGCCCCAGCGGAAGCACAGCTCAG
AGCTGGTCTGCCATGGGACATCCTGGTCCCCTCCTGCAGCTGCTGGTGTGCTTCTTACCCTGCC
CCTGCACCTCATGGCTCTGCTGGGCTGCTGGCAGCCCCTGTGCAAAGCTACTTCCCCTACCTGA
TGGCCGTGCTGACTCCCAAGAGCAACCGCAAGATGGAGAGCAAGAAACGGGAGCTCTTCAGCCAG
ATAAAGGGGCTTACAGGAGCCTCCGGGAAAGTGGCCCTACTGGAGCTGGGCTGCGGAACCGGAGC
CAACTTTCAGTTCTACCCACCGGGCTGCAGGGTCACCTGCCTAGACCCAAATCCCCACTTTGAGA
AGTTCCTGACAAAGAGCATGGCTGAGAACAGGCACCTCCAATATGAGCGGTTTGTGGTGGCTCCT
GGAGAGGACATGAGACAGCTGGCTGATGGCTCCATGGATGTGGTGGTCTGCACTCTGGTGTG
CTCTGTGCAGAGCCCAAGGAAGGTCTGCAGGAGGTCCGGAGAGTACTGAGACCGGGAGGTGTGC
TCTTTTTCTGGGAGCATGTGGCAGAACCATATGGAAGCTGGGCCTTCATGTGGCAGCAAGTTTTC
GAGCCCACCTGGAAACACATTGGGGATGGCTGCTGCCTCACCAGAGAGACCTGGAAGGATCTTGA
GAACGCCAGTTCTCCGAAATCCAATGGAACGACAGCCCCCTCCCTTGAAGTGGCTACCTGTTG
GGCCCCACATCATGGGAAAGGCTGTCAAACAATCTTTCCCAAGCTCCAAGGCACTCATTGCTCC
TTCCCAGCCTCCAATTAGAACAGCCACCCACCAGCCTATCTATCTTCCACTGAGAGGGACCTA
GCAGAATGAGAGAAGACATTCATGTACCACCTACTAGTCCCTCTCTCCCCAACCTCTGCCAGGGC
AATCTCTAACTTCAATCCCGCCTTCGACAGTGAAAAAGCTCTACTTCTACGCTGACCCAGGGAGG
AAACACTAGGACCCTGTTGTATCCTCAACTGCAAGTTTCTGGACTAGTCTCCAACGTTTGCCTC
CCAATGTTGTCCCTTTCCTTCGTTCCCATGGTAAAGCTCCTCTCGCTTTCCTCCTGAGGCTACAC
CCATGCGTCTCTAGGAACTGGTCACAAAAGTCATGGTGCCTGCATCCCTGCCAAGCCCCCTGAC
CCTCTCTCCCCACTACCACCTTCTTCTGAGCTGGGGGCACCAGGGAGAATCAGAGATGCTGGGG
ATGCCAGAGCAAGACTCAAAGAGGCAGAGTTTTGTTCTCAAATATTTTTTAATAAATAGACGAA
ACCACG

FIGURE 28

MDILVPLLQLLVLLLTPLPHLMALLGCWQPLCKSYFPYLMVLTPKSNRKMESKKRELFSSQIKGL
TGASGKVALLELGCCTGANFQFYPGCRVTCCLDPNPHFEKFLTKSMAENRHLQYERFVVAPGEDM
RQLADGSMDDVVCTLVLCVSVQSPRKVLQEVRRVLRPGGVLFWEHVAEYPYGSWAFMWQQVFPEPTW
KHIGDGCCLTRETWKDLENAQFSEIQMERQPPPLKWLPGVPHIMGKAVKQSFSSKALICSFPSL
QLEQATHQPIYLLPLRGT

Important features:

Signal peptide:

amino acids 1-23

Leucine zipper pattern.

amino acids 10-32

N-myristoylation sites.

amino acids 64-70, 78-84, 80-86, 91-97, 201-207

FIGURE 29

CAATGTTTGCCATCCACCTCCCCAAGCCCCTTACCTATGCTGCTGCTAACGCTGCTGCTGCT
GCTGCTGCTGCTTAAAGGCTCATGCTTGGAGTGGGGACTGGTCGGTGCCAGAAAGTCTCTTCTG
CCACTGACGCCCCCATCAGGGATTGGGCCTTCTTTCCCCCTTCCTTTCTGTGTCTCCTGCCTCAT
CGGCCTGCCATGACCTGCAGCCAAGCCCAGCCCCGTGGGGAAGGGGAGAAAGTGGGGGATGGCTTA
AGAAAGCTGGGAGATAGGGAACAGAAGAGGGTAGTGGGTGGGCTAGGGGGCTGCCTTATTTAAA
GTGGTTGTTTATGATTCTTATACTAATTTATACAAAGATATTAAGGCCCTGTTTATTAAAGAAATT
GTTCCCTTCCCCTGTGTTCAATGTTTGTAAGATTGTTCTGTGTAATATGTCTTTATAATAAAC
AGTTAAAAGCTGAAAAAAAAAAAAAAAAAAAAAAAAAAAA

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500

FIGURE 31

GTTTGAATTCCTTCAACTATACCCACAGTCCAAAAGCAGACTCACTGTGTCCCAGGCTACCAAGTT
CCTCCAAGCAAGTCATTTCCCTTATTTAACCGATGTGTCCCTCAAACACCTGAGTGCTACTCCCT
ATTTGCATCTGTTTTGATAAATGATGTTGACACCCTCCACCGAATTCCTAAGTGGAAATCATGTCGG
GAAGAGATACAATCCTTGGCCTGTGTATCCTCGCATTAGCCTTGTCTTTGGCCATGATGTTTACC
TTCAGATTCATCACCACCCTTCTGGTTCACATTTTCATTTTCATTGGTTATTTTGGGATTGTTGTT
TGTCTGCGGTGTTTTATGGTGGCTGTATTATGACTATAACCAACGACCTCAGCATAGAATTGGACA
CAGAAAGGGAAAATATGAAGTGCCTGCTGGGGTTTGCTATCGTATCCACAGGCATCACGGCAGTG
CTGCTCGTCTTGATTTTTGTTCTCAGAAAGAGAATAAAATTGACAGTTGAGCTTTTCCAAATCAC
AAATAAAGCCATCAGCAGTGCTCCCTTCTGCTGTTCCAGCCACTGTGGACATTTGCCATCCTCA
TTTTCTTCTGGGTCTCTGGGTGGCTGTGCTGCTGAGCCTGGGAACTGCAGGAGCTGCCAGGTT
ATGGAAGGCGGCCAAGTGAATATAAGCCCTTTTCGGGCATTTCGGTACATGTGGTTCGTACCATTT
AATTGGCCTCATCTGGACTAGTGAATTCATCCTTGCCTGCCAGCAAATGACTATAGCTGGGGCAG
TGGTTACTTGTATTTCAACAGAAGTAAAAATGATCCTCCTGATCATCCCATCCTTTTCGTCTCTC
TCCATTCTCTTCTTACCATCAAGGAACCGTTGTGAAAGGTCATTTTTAATCTCTGTGGTGAG
GATTCCGAGAATCATTGTCATGTACATGCAAACGCACTGAAAGAACAGCAGCATGGTGCATGT
CCAGGTACCTGTTCCGATGCTGCTACTGCTGTTTCTGGTGTCTTGACAAATACCTGCTCCATCTC
AACCAGAATGCATATACTACAACCTGCTATTAATGGGACAGATTTCTGTACATCAGCAAAAGATGC
ATTCAAAATCTTGTCCAAGAACTCAAGTCACTTTACATCTATTAAGTCTTTGGAGACTTCATAA
TTTTTCTAGGAAAGGTGTAGTGGTGTGTTTCACTGTTTTTGGAGGACTCATGGCTTTTAACTAC
AATCGGGCATTCAGGTGTGGGCAGTCCCTCTGTTATTGGTAGCTTTTTTTGCCTACTTAGTAGC
CCATAGTTTTTTATCTGTGTTTGAAGTGTGCTGGATGCACTTTCCTGTGTTTTGCTGTTGATC
TGGAACAAATGATGGATCGTCAGAAAAGCCCTACTTTATGGATCAAGAATTTCTGAGTTTCGTA
AAAAGGAGCAACAAATTAACAATGCAAGGGCACAGCAGGACAAGCACTCATTAAGGAATGAGGA
GGGAACAGAACTCCAGGCCATTGTGAGATTAGATACCCATTTAGGTATCTGTACCTGGAAAACATT
TCCTTCTAAGAGCCATTTACAGAATAGAAGATGAGACCACTAGAGAAAAGTTAGTGAATTTTTTT
TTAAAGACCTAATAACCCCTATTCTCTCCTCAAAA

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

FIGURE 32

MSGRDTILGLCILALALSLAMMFTFRFITLLVHIFISLVILGLLFVCGVLWWLYDYDNDLSIE
LDTARENMKCVLGFVAVSTGITAVLLVLI FVLRKRRIKLTVELFQITNKAISSAPFLLFQPLWTF
ILIFFWVWVAVLLSLGTAGAAQVMEGGQVEYKPLSGIRYMWSYHLIGLIWTSEFILACQOMTIA
GAVVTCYFNRSKNDPPDHPILSSLSILFFYHQGTVVKGSFLISVVRIPRIIVMYMQLALKEQQHG
ALSRYLFRCCYCCFWCLDKYLLHLNQNAYTTTAINGTDFCTSAKDAFKILSKNSSHFTSINCFGD
FIIFLGKVLVVCFTVFGGLMAFNYNRAQVWAVPLLLVAFFAYLVVHSFLSVFETVLDALFLCFA
VDLETNDGSSEKPYFMDQEFLSFVKRSNKLNNARAQQDKHSRNEEGTELOAIVR

Important features:

Signal peptide:

amino acids 1-20

Putative transmembrane domains:

amino acids 35-54, 75-97, 126-146, 185-204, 333-350, 352-371

N-glycosylation sites.

amino acids 204-208, 295-299, 313-317

N-myristoylation sites.

amino acids 147-153, 178-184, 196-202, 296-275, 342-348

FIGURE 34

MRTVVLTMKASVIEMFLVLLVTVGVHSNKETAKKIKRPFVTPQINCDVKAGKIIDPEFIVKCPAG
CQDPKYHVYGTDVYASYSSVCGAAVHSGVLDNSGGKILVRKVAGQSGYKGSYSNGVQSLSLPRWR
ESFIVLESKPKKGVTPSALTYSSSKSPAAQAGETTKAYQRPPIPGTTAQPVTLMQLLAVTVAVA
TPPTLPRPSPAASTTIPRPQSVGHRSEQEMDLWSTATYTSSQNRPRADPGIQRQDPSGAAFOKP
VGADVSLGLVPKEELSTQSLEPVS LGDPNCKIDLSFLIDGSTS IGKRRFRIQKQLLADVAQALDI
GPAGPLMGVVQYGDNPATHFNLKHTNSRDLKTAIEKITQRGGLSNVGRAISFVTKNFFSKANGN
RSGAPNVVVVMVDGWPTDKVEEASRLARESGINIFFITIEGAAENEKQYVVEPNFANKAVCRTNG
FYSLHVQSWFGLHKTLPVLRKVCDTDLACSKTCLNSADIGFVIDGSSSVGTGNFRTVLQFVTN
LTKEFEISDTRIGAVQYTYEQRLEFGFDKYSSKPDILNAIKRVGYWSSGGTSTGAAINFALEQL
FKKSKPNKRKLMILITDGRSYDDVRIPAMAAHLKGVITYAIGVAAQAEELEVIATHPARDSFF
VDEFDNLHQYVPRIIQNICTEFNSQPRN

Important features:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 181-200

N-glycosylation sites.

amino acids 390-394, 520-524

N-myristoylation sites.

amino acids 23-29, 93-99, 115-121, 262-268, 367-373, 389-395,
431-437, 466-472, 509-515, 570-576, 571-577, 575-581, 627-633

Amidation site.

amino acids 304-308

FIGURE 35

CCGAGCACAGGAGATTGCCTGCGTTTTAGGAGGTGGCTGCGTTGTGGGAAAAGCTATCAAGGAAGAAATTGC
CAAACCATGTCTTTTTTCTGTTTTAGAGTAGTTCACAACAGATCTGAGTGTTTAATTAAGCATGGAAT
ACAGAAAACAACAAAAAATAAGCTTTAATTTTCATCTGGAATCCACAGTTTTCTTAGCTCCCTGGACCC
GGTTGACCTGTTGGCTCTTCCCCTGGCTGCTCTATCACGTGGTGTCTCCGACTACTCACCCCGAGTGTA
AAGAACCTTCGGCTCGCGTCTCTGAGCTGCTGTGGATGGCCTCGGCTCTCTGGACTGTCTTCCGAGTA
GGATGTCAGTGAATCCCTCAAATGGAGCCTCTGCTGCTGTCAGTCTGAGTTTCTTTGTGATGTGGTAC
CTCAGCCTTCCCCACTACAATGTGATAGAACCGTGAAGTGGATGTACTTCTATGAGTATGAGCCGATTA
CAGACAAGACTTTCACCTTCACACTTCGAGAGCATTCAAAGTCTCTCATCAAATCCATTTCTGGTCATTC
TGGTACCTCCCACCCTTCAGATGTGAAAGCCAGGCAGGCCATTAGAGTTACTTGGGGTGAAAAAAGTCT
TGGTGGGATATGAGTTCTTACATTTTTCTTATTAGGCCAAGAGGCTGAAAAGGAAGACAAAATGTTGGC
ATTGTCCTTAGAGGATGAACACCTTCTTTATGGTACATAATCCGACAAGATTTTTTAGACACATATAATA
ACCTGACCTTAAAACCATTATGGCATTGAGTGGGTAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT
AAGACAGACACTGATGTTTTTCACTCAATACTGGCAATTTAGTGAAGTATCTTTTAAACCTAAACCACTCAGA
GAAGTTTTTACAGGTTATCCTCTAATGATAATTATCCTATAGAGATTTTACAAAAAACCCATATTT
CTTACCAGGAGTATCCTTTCAAGGTGTTCCCTCCATACTGCAGTGGGTGGGTTATATAATGTCAGAGAT
TTGGTGCCAAGGATCTATGAAATGATGGGTACGTAACCCATCAAGTTTGAAGATGTTTATGTCGGGAT
CTGTTTGAATTTATTAAGTGAACATTCATATTCAGAAGACAAAATCTTTTCTTTCTATATAGAATCC
ATTTGGATGTCTGTCAACTGAGACGTGTGATTGCAGCCATGGCTTTTCTTCCAAGGAGATCATCACTTTT
TGGCAGGTGATGCTAAGGAACACCACATGCCATTATTAAGTTCACATTCTACAAAAAGCCTAGAAGGACAG
GATACCTTGTGAAAGTGTAAATAAAGTAGTACTGTGAAAATTCATGGGAGGTGAGTGTGCTGGCTT
ACACTGAACTGAACTCATGAAAAACCCAGACTGGAGACTGGAGGTTACACTTGTGATTTATTAGTCAGG
CCCTTCAAAGATGATATGTGGAGGAATTAATATAAAGGAATTGGAGGTTTTTGCTAAAAGAAATTAATAGG
ACCAAACAATTTGGACATGTCAATCTGTAGACTAGAATTTCTTAAAAGGGTGTACTGAGTTATAAGCTCA
CTAGGCTGTA AAAACAAAACAATGTAGAGTTTTATTATTGAACAATGTAGTCACTTGAAGGTTTTGTGTA
TATCTTATGTGATTACCAATTTAAAAATATATGTAGTCTGTGTCAAAAAACTTCTCACTGAAGTTATA
CTGAACAAAATTTACCTGTTTTTGGTCATTTATAAAGTACTTCAAGATGTTGCAGTATTTACAGTTATT
ATTATTTAAAATTACTTCAACTTTGTGTTTTTAAATGTTTTGACGATTTCAATACAAGATAAAAAGGATAG
TGAATCATCTTTACATGCAACATTTTCCAGTACTTAACTGATCAGTTTATATATGATACATCACTCCA
TTAATGTAAAGTCATAGGTCATTATTGCATATCAGTAATCTCTTGGACTTGTAAATATTTTACTGTGGT
AATATAGAGAAGAATTAAGCAAGAAAATCTGAAA

FIGURE 36

MASALWTVLPSRMSLRSLKWSLLLLSLLSFFVMWYLSLPHYNVIERVNWMYFYEYEPYRQDFHF
TLREHSNCSHQNPFLVILVTSHPSDVKARQAIRVTWGEKKSWWGYEVLTFLLGQEAEKEDKMLA
LSLEDEHLLYGDIIRQDFLDTYNNLTLKTIMAFRWVTEFCPNAKYVMKTDTDVDFINTGNLVKYL
NLNHSEKFFTGYPLIDNYSYRGFYQKTHISYQYEPFKVFPYCSGLGYIMSRDLVPRIYEMMGHV
KPIKFEDVYVIGICLNLLKVNIIHPEDTNLFFLYRIHLDVCQLRRVIAAHGFSSKEIITFWQVMLR
NTTCHY

Important features:

Type II transmembrane domain:

amino acids 20-39

N-glycosylation sites.

amino acids 72-76, 154-158, 198-202, 212-216, 326-330

Glycosaminoglycan attachment site.

amino acids 239-243

Ly-6 / u-PAR domain proteins.

amino acids 23-37

N-myristoylation site.

amino acids 271-277

FIGURE 37

CGCTCGGGCACCAGCCGCGCAAGGATGGAGCTGGGTTGCTGGACGAGTTGGGGCTCACTTTTCTTCAGCTCCTTCTCATC
TCGTCTTGCCAAGAGAGTACACAGTCATTAATGAAGCCTGCCCTGGAGCAGAGTGGAAATATCATGTGTGGGAGTCTGTG
AATATGATCAGATTGAGTGCCTTGCCCCGAAAGAGGGAAGTCGTGGTTATACCATCCCTTGCTGCAGGAATGAGGAGAA
TGAGTGTGACTCCTGCCTGATCCACCAGGTTGTACCATCTTTGAAAACGCAAGAGCTGCCGAAATGGCTCATGGGGGGT
ACCTTGATGACTTCTATGTGAAGGGTTCTACTGTGCAGAGTGCCGAGCAGGCTGGTACGGAGGAGACTGCATGCCATGTG
GCCAGTTCTGCAGCCCCAAGGGTCAGATTTGTGAAAGCTATCCCTAAATGCTCACTGTGAATGGACCATTTCATGC
TAAACCTGGGTTTGTATCCAACTAAGATTTGTGATGTTGAGTCTGGAGTTGACTACATGTGCCAGTATGACTATGTTGAG
GTTGATGAGAGACAACCGCATGGCCAGATCATCAAGCGTGTCTGTGGCAACGAGCGGCCAGCTCCTATCCAGAGCATAG
GATCCTCACGCCGCTCCTTCCACTCCGATGGCTCCAAGAATTTGACGGTTTCCATGCCATTTATGAGGAGATCACAGC
ATGCTCCTCATCCCTTGTTCATGACGGCAGTGCGTCTTGTACAAGGCTGGATCTTACAAGTGTGCCTGCTGGCAGGC
TATACTGGGACGGCTGTGAAAATCCTTGAAGAAAAGAACTGCTCAGACCTGGGGGCCAGTCAATGGGTACCAGAAAA
TAACAGGGGGCCCTGGGCTTATCAACGGAGCCATGTAAAATGGCACCGTGGTGTCTTCTTTTGAACACTCCTATGT
TCTTAGTGGCAATGAGAAAAGAACTTCCAGCAGAAATGGAGAGTGGTCAGGGAACAGCCCATCTGCATAAAAAGCTGCGGA
GAACCAAAGATTTGAGACTGGTGAGAGAGAGTCTTCCGATGCAGGTTGAGTCAAGGAGACACCATTACACCAGCTAT
ACTCAGCGGCTTCAGCAAGCAGAACTGCAGAGTGCCCTACCAAGAAGCCAGCCCTTCCCTTTGGAGATCTGCCATGGG
ATACCAACATCTGCATACCCAGCTCCAGTATGAGTGCATCTCACCTTCTACCGCCCTGGGAGCAGCAGGAGGACATGT
CTGAGGACTGGGAAGTGGATGGCGGGCCACCTCCGATCCCTATCTGCGGGAAAATGAGAACATCACTGCTCCAAGA
CCCAGGGTTGCGCTGGCCGTGGCAGGCAGCCATCTACAGGAGACCAGCGGGTGCATGACGGCAGCCTACACAAGGGAGC
GTGGTTCTTAGTCTGCAGCGGTGCCCTGGTGAATGAGCGCACTGTGGTGGTGGCTGCCACTGTGTTACTGACCTGGGGAG
GTCACCATGATCAAGACAGCAGACCTGAAAGTGTGTTTGGGAAAATCTACCGGATGATGACCGGGATGAGAAGACCATCC
AGAGCCTACAGATTTCTGCTATCATTCTGCATCCCAACTATGACCCCATCTGCTTGTATGTGACATCGCCATCCTGAAGCT
CCTAGACAAGGCCGATCAGCACCCGAGTCCAGCCATCTGCCTCGCTGCCAGTGGGATCTCAGCACTTCTTCCAGGAG
TCCCACATCACTGTGGCTGGCTGGAATGTCTGGCAGACGTGAGGAGCCCTGGTTCAAGAACGACACACTGCGCTTGGGG
TGGTCAGTGTGGTGGACTCGCTGCTGTGTGAGGAGCAGCATGAGGACCATGGCATCCAGTGAAGTGTCACTGATAACATGTT
CTGTGCCAGCTGGGAACCCACTGCCCTTCTGATATCTGCAGTGCAGAGCAGGAGGATCGCGGCTGTGCTCTCCGGGA
CGAGCATCTCTGAGCCAGCTGGCATCTGATGGGACTGGTCACTGAGTATGATAAAACATGCAGCCACAGGCTCTCCA
CTGCTTCACCAAGTGTGCTGCTTTAAAGACTGGATGAAAGAAATGAAATGAACCATGCTCATGCACTCCTTGAGAAG
TGTTCTGTATATCCGCTGTACGTGTGTCATTGCGTGAAGCAGTGTGGGCTGAAGTGTGATTTGGCCTGTGAAGTGGCT
GTGCCAGGGCTTCTGACTTCAGGGACAAAACCTCAGTGAAGGAGTGAAGTACCTCCATGCTGGTAGGCTGATGCCGCTCCA
CTACTAGGACAGCCAATGGAAGATGCCAGGCTTGCAGAAGTAAGTTCTTCAAAGAAGACCATATACAAAACCTCTCCA
CTCCACTGACCTGGTGGTCTTCCCAACTTTCAGTTATACGAAATGCCATCAGCTTGACCAGGGAAGATCTGGGCTTCATGAG
GCCCCTTTGAGGCTCTCAAGTCTAGAGAGCTGCTGTGGGACAGCCAGGGCAGCAGAGCTGGGATGTGGTGCATGCTT
TGTGTACATGGCCACAGTACAGTCTGGTCTTTTCTTCCCATCTCTTGTACACATTTAATAAAAATAAGGGTGGCTTCT
GAAC TACAA
AA

CGCTCGGGCACCAGCCGCGCAAGGATGGAGCTGGGTTGCTGGACGAGTTGGGGCTCACTTTTCTTCAGCTCCTTCTCATC
TCGTCTTGCCAAGAGAGTACACAGTCATTAATGAAGCCTGCCCTGGAGCAGAGTGGAAATATCATGTGTGGGAGTCTGTG
AATATGATCAGATTGAGTGCCTTGCCCCGAAAGAGGGAAGTCGTGGTTATACCATCCCTTGCTGCAGGAATGAGGAGAA
TGAGTGTGACTCCTGCCTGATCCACCAGGTTGTACCATCTTTGAAAACGCAAGAGCTGCCGAAATGGCTCATGGGGGGT
ACCTTGATGACTTCTATGTGAAGGGTTCTACTGTGCAGAGTGCCGAGCAGGCTGGTACGGAGGAGACTGCATGCCATGTG
GCCAGTTCTGCAGCCCCAAGGGTCAGATTTGTGAAAGCTATCCCTAAATGCTCACTGTGAATGGACCATTTCATGC
TAAACCTGGGTTTGTATCCAACTAAGATTTGTGATGTTGAGTCTGGAGTTGACTACATGTGCCAGTATGACTATGTTGAG
GTTGATGAGAGACAACCGCATGGCCAGATCATCAAGCGTGTCTGTGGCAACGAGCGGCCAGCTCCTATCCAGAGCATAG
GATCCTCACGCCGCTCCTTCCACTCCGATGGCTCCAAGAATTTGACGGTTTCCATGCCATTTATGAGGAGATCACAGC
ATGCTCCTCATCCCTTGTTCATGACGGCAGTGCGTCTTGTACAAGGCTGGATCTTACAAGTGTGCCTGCTGGCAGGC
TATACTGGGACGGCTGTGAAAATCCTTGAAGAAAAGAACTGCTCAGACCTGGGGGCCAGTCAATGGGTACCAGAAAA
TAACAGGGGGCCCTGGGCTTATCAACGGAGCCATGTAAAATGGCACCGTGGTGTCTTCTTTTGAACACTCCTATGT
TCTTAGTGGCAATGAGAAAAGAACTTCCAGCAGAAATGGAGAGTGGTCAGGGAACAGCCCATCTGCATAAAAAGCTGCGGA
GAACCAAAGATTTGAGACTGGTGAGAGAGAGTCTTCCGATGCAGGTTGAGTCAAGGAGACACCATTACACCAGCTAT
ACTCAGCGGCTTCAGCAAGCAGAACTGCAGAGTGCCCTACCAAGAAGCCAGCCCTTCCCTTTGGAGATCTGCCATGGG
ATACCAACATCTGCATACCCAGCTCCAGTATGAGTGCATCTCACCTTCTACCGCCCTGGGAGCAGCAGGAGGACATGT
CTGAGGACTGGGAAGTGGATGGCGGGCCACCTCCGATCCCTATCTGCGGGAAAATGAGAACATCACTGCTCCAAGA
CCCAGGGTTGCGCTGGCCGTGGCAGGCAGCCATCTACAGGAGACCAGCGGGTGCATGACGGCAGCCTACACAAGGGAGC
GTGGTTCTTAGTCTGCAGCGGTGCCCTGGTGAATGAGCGCACTGTGGTGGTGGCTGCCACTGTGTTACTGACCTGGGGAG
GTCACCATGATCAAGACAGCAGACCTGAAAGTGTGTTTGGGAAAATCTACCGGATGATGACCGGGATGAGAAGACCATCC
AGAGCCTACAGATTTCTGCTATCATTCTGCATCCCAACTATGACCCCATCTGCTTGTATGTGACATCGCCATCCTGAAGCT
CCTAGACAAGGCCGATCAGCACCCGAGTCCAGCCATCTGCCTCGCTGCCAGTGGGATCTCAGCACTTCTTCCAGGAG
TCCCACATCACTGTGGCTGGCTGGAATGTCTGGCAGACGTGAGGAGCCCTGGTTCAAGAACGACACACTGCGCTTGGGG
TGGTCAGTGTGGTGGACTCGCTGCTGTGTGAGGAGCAGCATGAGGACCATGGCATCCAGTGAAGTGTCACTGATAACATGTT
CTGTGCCAGCTGGGAACCCACTGCCCTTCTGATATCTGCAGTGCAGAGCAGGAGGATCGCGGCTGTGCTCTCCGGGA
CGAGCATCTCTGAGCCAGCTGGCATCTGATGGGACTGGTCACTGAGTATGATAAAACATGCAGCCACAGGCTCTCCA
CTGCTTCACCAAGTGTGCTGCTTTAAAGACTGGATGAAAGAAATGAAATGAACCATGCTCATGCACTCCTTGAGAAG
TGTTCTGTATATCCGCTGTACGTGTGTCATTGCGTGAAGCAGTGTGGGCTGAAGTGTGATTTGGCCTGTGAAGTGGCT
GTGCCAGGGCTTCTGACTTCAGGGACAAAACCTCAGTGAAGGAGTGAAGTACCTCCATGCTGGTAGGCTGATGCCGCTCCA
CTACTAGGACAGCCAATGGAAGATGCCAGGCTTGCAGAAGTAAGTTCTTCAAAGAAGACCATATACAAAACCTCTCCA
CTCCACTGACCTGGTGGTCTTCCCAACTTTCAGTTATACGAAATGCCATCAGCTTGACCAGGGAAGATCTGGGCTTCATGAG
GCCCCTTTGAGGCTCTCAAGTCTAGAGAGCTGCTGTGGGACAGCCAGGGCAGCAGAGCTGGGATGTGGTGCATGCTT
TGTGTACATGGCCACAGTACAGTCTGGTCTTTTCTTCCCATCTCTTGTACACATTTAATAAAAATAAGGGTGGCTTCT
GAAC TACAA
AA

FIGURE 38

MELGCWTQLGLTFLQLLLISSLPREYTVINEACPGAENIMCRECCEYDQIECVCPGKREVVGYT
IPCCRNEENECDSCLIHGCTIFENCKSCRNGSWGGLDDFYVKGFYCAECRAGWYGGDCMRCGQ
VLRAPKGQILLESYPLNAHCEWTIHAKPGFVIQLRFVMLSLEFDYMCQYDYVEVRDGDNRDQII
KRVCGNERPAPIQSIGSSLVLFHSDGSKNFDGFHAIYEEITACSSSPCFHDGTCVLDKAGSYKC
ACLAGYTGQRCEENLEERNCSDPGGPVNGYQKITGGPGLINGRHAKIGTVVSFFCNNSYVLSGNE
KRTCQONGEWSGKQPICIKACREPKISDLVRRRVLPMQVQSRETPHQLYSAAFSKQKLSAPTK
KPALPFGDLPMGYOHLHTQLQYECISPFYRRLGSSRRTCLRTGKWSGRAPSCIPICGIENITAP
KTQGLRWPQAAIYRRTSGVHDGSLHKGAWFLVCSGALVNERTVVVAHCVTDLGKVTMIKTADL
KVVLGKFYRDDDRDEKTIQSLQISAIILHPNYDPILLDADIAILKLLDKARISTRVQPICLAASR
DLSTSFQESHITVAGWNVLADVRSPGFKNDTLRSGVVSVDLLCEEQHEDHGIPVSVTDNMFCA
SWEPTAPSDICTAETGGIAAVSFPGRASPEPRWHLMGLVSWSYDKTCSHRLSTAFTKVLPEFKDWI
ERNMK

Important features of the protein:

Signal peptide:

amino acids 1-23

EGF-like domain cysteine pattern signature.

amino acids 260-272

N-glycosylation sites.

amino acids 96-100, 279-283, 316-320, 451-455, 614-618

N-myristoylation sites.

amino acids 35-41, 97-103, 256-262, 284-290, 298-304, 308-314,
474-480, 491-497, 638-644, 666-672

Amidation site.

amino acids 56-60

Serine proteases, trypsin family.

amino acids 489-506

CUB domain proteins profile.

amino acids 150-167

FIGURE 40

MKALLLLVLPWLS PANYIDNVGNLHFLYSELCKGASHYGLTKDRKRRSQDGCPCDGCASLTATAPS
PEVSAATISLMTDEPGLDNPAYVSSAEDGQPAISPVDSGRSNRTRARPFERSTIRSRSEFKKINR
ALSVLRRTKSGSAVANHADQGRENSENTAPEVFPRLYHLIPDGEITSIKINRVDPSELSIRLV
GGSETPLVHIIQHIYRDGVIARDGRLLPGDIIKLVNGMDISNVPHNYAVRLLRQPCQVLWLTVM
REQKFRSRNNGQAPDAYRPRDDSFHVILNKSSPEEQLGIKLVKRVDEPGVFIENVLDGGVAYRHG
QLEENDRVLAINGHDLRYGSPESAHLIQASERRVHLVVSQRVQRSPDIFQEAGWNSNGSWSPG
PGRSNTPKPLHPTITCHEKVVNIQKDPGESLGMTVAGGASHREWDLPIYVISVEPGGVISRDGR
IKTGDILLNVGVELTEVSRSEAVALLKRTSSSIVLKALEVKEYEPQEDCSSPAALDSNHNMAPP
SDWSPSWVMWLELPRCLYNCKDIVLRRNTAGSLGFCIVGGYEEYNGNKPFFIKSIVEGTPAYNDG
RIRCGDILLAVNGRSTSGMIHAACLARLLKELKGRITLTIVSWPGTFL

Important features:

Signal peptide:

amino acids 1-15

N-glycosylation sites.

amino acids 108-112, 157-161, 289-293, 384-388

Tyrosine kinase phosphorylation sites.

amino acids 433-441, 492-500

N-myristoylation sites.

amino acids 51-57, 141-147, 233-239, 344-350, 423-429, 447-453,
467-473, 603-609

FIGURE 41

ACCAGGCATTGTATCTTCAGTTGTCATCAAGTTCGCAATCAGATTGGAAAAGCTCAACTTGAAGCTTT
CTTGCCTGCAGTGAAGCAGAGAGATAGATATTATTACGTAATAAAAAACATGGGCTTCAACCTGA
TTCCACCTTTCCCTACAAATCCGATTACTGTTGCTGTTGACTTTGTGCCTGACAGTGGTTGGGTGGGC
CACCAGTAACTACTTCGTGGGTGCCATTCAAGAGATTCCCTAAAGCAAAGGAGTTCATGGCTAATTTCC
ATAAGACCCTCATTTTGGGGAAGGGAAAACTCTGACTAATGAAGCATCCACGAAGAAGGTAGAACTT
GACAACGTCTCTTCTGTCTCCTTACCTCAGAGGCCAGAGCAAGCTCATTTTCAAACCAGATCTCAC
TTTGGAAAGAGGTACAGGCAGAAAATCCCAAAGTGTCCAGAGGCCGGTATCGCCCTCAGGAATGTAAG
CTTTACAGAGGGTCCGCATCTCGTTCCCCACCGGAACAGAGAGAAACACCTGATGTACCTGCTGGAA
CATCTGCATCCCTTCTGCAGAGGCAGCAGCTGGATTATGGCATCTACGTCATCCACCAGGCTGAAGG
TAAAAGTTTAATCGAGCCAAACTCTTGAATGTGGGCTATCTAGAAGCCCTCAAGGAAGAAAATTGGG
ACTGCTTTATATCCACGATGTGGACCTGGTACCCGAGAATGACTTTAACCTTTACAAGTGTGAGGAG
CATCCCAAGCATCTGGTGGTTGGCAGGAACAGCACTGGGTACAGGTTACGTTACAGTGGATATTTTGG
GGGTGTTACTGCCCTAAGCAGAGAGCAGTTTTTCAAGGTGAATGGATTCTCTAACAACTACTGGGGAT
GGGGAGCGAAGACGATGACCTCAGACTCAGGGTTGAGCTCCAAAGAATGAAAATTTCCCGGCCCTG
CCTGAAGTGGGTAAATATAACAATGGTCTTCCACACTAGAGACAAAGGCAATGAGGTGAACGCAGAACG
GATGAAGCTCTTACACCAAGTGTCCAGAGTCTGGAGAACAGATGGGTTGAGTAGTTGTTCTTATAAAT
TAGTATCTGTGGAACACAATCCTTTATATATCAACATCACAGTGGATTTCTGGTTTGGTGCATGACCC
TGGATCTTTGGTGATGTTTGGGAAGAACTGATCTTTGTTTGAATAATTTTGGCCTAGAGACTTCAA
ATAGTAGCACACATTAAGAACCTGTTACAGCTCATTGTTGAGCTGAATTTTCTTTTTGTATTTTCT
TAGCAGAGCTCCTGGTGATGTAGAGTATAAAACAGTTGTAACAAGACAGCTTTCTTAGTCATTTTGAT
CATGAGGGTTAAATATTGTAATATGGATACTTGAAGGACTTTATATAAAAGGATGACTCAAAGGATAA
AATGAACGCTATTTGAGGACTCTGGTTGAAGGAGATTTATTTAAATTTGAAGTAATATATATGGGAT
AAAAGGCCACAGGAAATAAGACTGCTGAATGTCTGAGAGAACCAGAGTTGTTCTCGTCCAAGGTAGAA
AGGTACGAAGATACAATACTGTTATTCATTTATCCTGTACAATCATCTGTGAAGTGGTGGTGTGAGGT
GAGAAGGCGTCCACAAAAGAGGGGAGAAAAGGCGACGAATCAGGACACAGTGAACCTGGGAATGAAGA
GGTAGCAGGAGGGTGGAGTGTCCGGCTGCAAAGGCAGCAGTAGCTGAGCTGGTTGCAGGTGCTGATAGC
CTTCAGGGGAGGACCTGCCAGGTATGCCTTCCAGTGTGCCCACCAGAGAATACATTTCTATTAGT
TTTTAAAGAGTTTTTGTAAAATGATTTTGTACAAGTAGGATATGAATTAGCAGTTTACAAGTTTACAT
ATTAATAATAATAAATATGTCTATCAAATACCTCTGTAGTAAAAATGTGAAAAGCAAAA

FIGURE 42

MGFNLT FHLSYKFRLLLLLTLCLTVVWATSNYFVGAIQEIPKAKEFMANFHKTLILGKGKLTN
EASTKKVELDNCPSVSPYLRGQSKLIFKPDLTLEEVQAENPKVSRGRYRPQECKALQRVAILVPH
RNREKHLMYLLEHLHPFLQRQQLDYGIYVIHQAEQKFNRAKLLNVGYLEALKEENWDCFI FHDV
DLVPENDFNLYKCEEHPKHLVVGGRNSTGYRLRYSGYFGGVTALSREQFFKVNFGFSNNYWGWWGED
DDLRLRVELQRMKISRPLPEVGKYTMVFHTRDKGNEVNAERMKLLHQVSRVWRDGLSSCSYKLV
SVEHNPLYINITVDFWFGA

Important features:

Signal peptide:

amino acids 1-27

N-glycosylation sites.

amino acids 4-8, 220-224, 335-339

Xylose isomerase proteins.

amino acids 191-202

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500

FIGURE 44

MALSSQIWAACL LLLLLLLASLTSGSVFPQQTGQLAELQPDRAGARASWMPMFQRRRRRDTHFPI
CIFCCGCCHRSKCGMCCCKT

Important features:

Signal peptide:

amino acids 1-24

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 58-59

N-myristoylation site.

amino acids 44-50

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 1-12

FIGURE 45

GTGGCTTCATTTTCAGTGGCTGACTTCCAGAGAGCAATATGGCTGGTTCCCAACATGCCTCACCC
TCATCTATATCCTTTGGCAGCTCACAGGGTCAGCAGCCTCTGGACCCGTGAAAGAGCTGGTCCGGT
TCCGTTGGTGGGGCCGTGACTTTCCCCCTGAAGTCCAAAGTAAAGCAAGTTGACTCTATTGTCTG
GACCTTCAACACAACCCCTCTTGTCCACATACAGCCAGAAGGGGGCACTATCATAGTGACCCAAA
ATCGTAATAGGGAGAGAGTAGACTTCCCAGATGGAGGCTACTCCCTGAAGCTCAGCAAAGTGAAG
AAGAATGACTCAGGGATCTACTATGTGGGGATATACAGCTCATCACTCCAGCAGCCCTCCACCCA
GGAGTACGTGCTGCATGTCTACGAGCACCTGTCAAAGCCTAAAGTCACCATGGGTCTGCAGAGCA
ATAAGAATGGCACCTGTGTGACCAATCTGACATGCTGCATGGAACATGGGGAAGAGGATGTGATT
TATACCTGGAAGGCCCTGGGGCAAGCAGCCAATGAGTCCCATAATGGGTCCATCCTCCCCATCTC
CTGGAGATGGGGAGAAAGTGATATGACCTTCATCTGCGTTGCCAGGAACCCTGTCCAGCAGAACT
TCTCAAGCCCCATCCTTGCCAGGAAGCTCTGTGAAGGTGCTGCTGATGACCCAGATTCCCTCCATG
GTCCTCCTGTGTCTCCTGTTGGTGCCCTCCTGCTCAGTCTCTTTGTACTGGGGCTATTTCTTTG
GTTTCTGAAGAGAGAGAGACAAGAAGAGTACATTGAAGAGAAGAAGAGAGTGGACATTTGTCGGG
AAACTCCTAACATATGCCCCATTCTGGAGAGAACACAGAGTACGACACAATCCCTCACACTAAT
AGAACAATCCTAAAGGAAGATCCAGCAAATACGGTTTACTCCACTGTGGAAATACCGAAAAAGAT
GGAAATCCCCACTCACTGCTCACGATGCCAGACACACCAAGGCTATTTGCCTATGAGAATGTTA
TCTTAGACAGCAGTGCACTCCCCTAAGTCTCTGCTCA

FIGURE 46

MAGSPTCLTLIYILWQLTGSAAAGPVKELVGSVGGAVTFPLKSKVKQVDSIVWTFNNTPLVTIQP
EGGTIIIVTQNRNRERVDFFPDGGYSLKLSKLRKNDSGIYYVGIYSSSLQQPSTQEYVLHVYEHLSK
PKVTMGLQSNKNGTCVTNLTCCMEHGEEVDIYTWKALGQAANESHNGSILPISWRWGESDMTFIC
VARNPVSARNFSSPILARKLCEGAADDPDSSMVLCLLLVPLLLSLFVLGLFLWFLKRERQEEYIE
EKKRVDICRETPNICPHSGENTEYDTIPHTNRTILKEDPANTVYSTVEIPKKMENPHSLLTMPDT
PRLFAYENVI

Important features:

Signal peptide:

amino acids 1-22

Transmembrane domain:

amino acids 224-250

Leucine zipper pattern.

amino acids 229-251

N-glycosylation sites.

amino acids 98-102, 142-146, 148-152, 172-176, 176-180, 204-208,
291-295

FIGURE 47

GGCTCGAGCGTTTTCTGAGCCAGGGGTGACCATGACCTGCTGCGAAGGATGGACATCCTGCAATGG
ATTCAGCCTGCTGGTTCTACTGCTGTTAGGAGTAGTTCTCAATGCGATACCTCTAATTGTCAGCT
TAGTTGAGGAAGACCAATTTTCTCAAACCCCATCTCTTGCTTTGAGTGGTGGTTCCCAGGAATT
ATAGGAGCAGGTCTGATGGCCATTCCAGCAACAACAATGTCCTTGACAGCAAGAAAAAGAGCGTG
CTGCAACAACAGAACTGGAATGTTTCTTTCATCATTTTTTCAGTGTGATCACAGTCATTGGTGCTC
TGTATTGCATGCTGATATCCATCCAGGCTCTCTTAAAAGGTCCTCTCATGTGTAATTCTCCAAGC
AACAGTAATGCCAATTGTGAATTTTCATTGAAAACATCAGTGACATTCCAGAATCCTTCAA
CTTGCAAGTGGTTTTTCAATGACTCTTGTGCACCTCCTACTGGTTTTCAATAAACCCACCAGTAACG
ACACCATGGCGAGTGGCTGGAGAGCATCTAGTTTCCACTTCGATTCTGAAGAAAACAAACATAGG
CTTATCCACTTCTCAGTATTTTTAGGTCTATTGCTTGTGGAAATCTGGAGGTCCTGTTGGGCT
CAGTCAGATAGTCATCGGTTTTCTTGGCTGTCTGTGTGGAGTCTCTAAGCGAAGAAGTCAAATTG
TGTAGTTTAAATGGGAATAAAATGTAAGTATCAGTAGTTTGAAAAAAAAA

FIGURE 48

MTCCEGWTSCNGFSLVLLVLLLVVLLGVLNLAIPLIVSLVEEDQFSQNPISCFEWWFPGIIGAGLMAIPA
TTMSLTARKRACCNNRTGMFLSSFFSVITVIGALYCLISIQALLKGPLMCNSPSNSNANCEFSL
KNISDIHPESFNLQWFFNDSCAPPTGFNKPTSNDTMASGWRASSFHFDFSEENKHRLIHFSVFLGL
LLVGILEVLFGLSQIVIGFLGCLCGVSKRRSQIV

Important features:

Transmembrane domains:

amino acids 10-31 (type II), 50-72, 87-110, 191-213

N-glycosylation sites.

amino acids 80-84, 132-136, 148-152, 163-167

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 223-227

N-myristoylation sites.

amino acids 22-28, 54-60, 83-89, 97-103, 216-222

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 207-218

TNFR/NGFR family cysteine-rich region protein.

amino acids 4-12

FIGURE 50

MERVTLALLLLAGLTALEANDPFANKDDPFYYDWKNLQLSGLICGGLLAIAGIAAVLSGKCKYKS
SQKQHSPVPEKAIPLITPGSATTC

Important features:

Signal peptide:

amino acids 1-16

Transmembrane domain:

amino acids 36-59

N-myristoylation sites.

amino acids 41-47, 45-51, 84-90

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7.

amino acids 54-67

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

FIGURE 51

GTGGACTCTGAGAAGCCCAGGCAGTTGAGGACAGGAGAGAGAAGGCTGCAGACCCAGAGGGAGGG
AGGACAGGGAGTCGGAAGGAGGAGGACAGAGGAGGGCACAGAGACGCAGAGCAAGGGCGCAAGG
AGGAGACCCTGGTGGGAGGAAGACTCTGGAGAGAGAGGGGGCTGGGCAGAGATGAAGTTCCAG
GGGCCCTGGCCTGCCTCCTGCTGGCCCTCTGCCTGGGCAGTGGGAGGCTGGCCCCCTGCAGAG
CGGAGAGGAAAGCACTGGGACAAATATTGGGGAGGCCCTTGGACATGGCCTGGGAGACGCCCTGA
GCGAAGGGGTGGGAAAGGCCATTGGCAAAGAGGCCGGAGGGGCAGCTGGCTCTAAAGTCAGTGAG
GCCCTTGGCCAAGGGACCAGAGAAGCAGTTGGCACTGGAGTCAGGCAGGTTCCAGGCTTTGGCGC
AGCAGATGCTTTGGGCAACAGGGTCGGGGAAGCAGCCATGCTCTGGGAAACTGGGCACGAGA
TTGGCAGACAGGCAGAAGATGTCAATTCGACACGGAGCAGATGCTGTCCGCGGCTCCTGGCAGGGG
GTGCTTGGCCACAGTGGTGTCTGGGAACTTCTGGAGGCCATGGCATCTTTGGCTCTCAAGGTGG
CCTTGGAGGCCAGGGCCAGGGCAATCCTGGAGGTCTGGGGACTCCGTGGGTCCACGGATACCCCG
GAAACTCAGCAGGCAGCTTTGGAATGAATCCTCAGGGAGCTCCCTGGGGTCAAGGAGGCAATGGA
GGGCCACCAAACCTTTGGGACCAACTCAGGGAGCTGTGGCCAGCCTGGCTATGGTTCAGTGAG
AGCCAGCAACCAGAATGAAGGGTGCACGAATCCCCACCATCTGGCTCAGGTGGAGGCTCCAGCA
ACTCTGGGGGAGGCAGCGGCTCACAGTCGGGCAGCAGTGGCAGTGGCAGCAATGGTGACAACAAC
AATGGCAGCAGCAGTGGTGGCAGCAGCAGTGGCAGCAGCAGTGGCAGCAGCAGTGGCGGCAGCAG
TGGCGGCAGCAGTGGTGGCAGCAGTGGCAACAGTGGTGGCAGCAGAGGTGACAGCGGCAGTGAGT
CCTCCTGGGGATCCAGCACCGGCTCCTCCTCCGGCAACCACGGTGGGAGCGCGGAGGAAATGGA
CATAAACCCGGGTGTGAAAAGCCAGGGAAATGAAGCCCGGGGAGCGGGGAATCTGGGATTCAGGG
CTTCAGAGGACAGGGAGTTCCAGCAACATGAGGGAAATAAGCAAAGAGGGCAATCGCCTCCTTG
GAGGCTCTGGAGACAATTATCGGGGGCAAGGGTTCGAGCTGGGGCAGTGGAGGAGGTGACGCTGTT
GGTGGAGTCAATACTGTGAACCTCTGAGACGTCTCCTGGGATGTTAACTTTGACACTTTCTGGAA
GAATTTTAAATCCAAGCTGGGTTTTCATCAACTGGGATGCCATAAAACAAGGACCAGAGAAGCTCTC
GCATCCCGTGACCTCCAGACAAGGAGCCACCAGATTGGATGGGAGCCCCCACACTCCCTCCTTAA
AACACCACCCTCTCATCACTAATCTCAGCCCTTGCCTTGAATAAACCTTAGCTGCCCCACAAA
AA
AA

FIGURE 52

MKFQGPLACLLLLALCLGSGEAGPLQSGEESTGTNIGEALGHGLGDALSEGKAIKKEAGGAAGSKVS
EALGQGTREAVGTGVRQVPGFGAADALGNRVGEEAHALGNTGHEIGRQAEDVIRHGADAVRGSWQGV
GHSGAWETSGGHGIFGSQGGGGQGNPGGLGTPVWHGYPGNSAGSFGMNPQGAPWQGGNGGPPNF
GTNTQGAQAQPGYGSVRASNQNEGCTNPPPSGSGGGSSNSGGGSGSQSGSSGSGSNGDNNNGSSSGGS
SSGSSSSSSSGSSGGSSGGSSGNSGGSRGDSGSESSWGSSTGSSSGNHGGSGGGNGHKPKGCEKPGNE
ARGSGESGIQGFRGQGVSSNMREISKEGNRLLGGSGDNRYRGSSWGSGGGDAVGGVNTVNSETSPGM
FNFDTFWKNFKSKLGFINWDAINKDQRSSRIP

Signal peptide:

amino acids 1-21

N-glycosylation site.

amino acids 265-269

Glycosaminoglycan attachment site.

amino acids 235-239, 237-241, 244-248, 255-259, 324-328, 388-392

Casein kinase II phosphorylation site.

amino acids 26-30, 109-113, 259-263, 300-304, 304-308

N-myristoylation site.

amino acids 17-23, 32-38, 42-48, 50-56, 60-66, 61-67, 64-70, 74-80,
90-96, 96-102, 130-136, 140-146, 149-155, 152-158, 155-161,
159-165, 163-169, 178-184, 190-196, 194-200, 199-205, 218-224,
236-242, 238-244, 239-245, 240-246, 245-251, 246-252, 249-252,
253-259, 256-262, 266-272, 270-276, 271-277, 275-281, 279-285,
283-289, 284-290, 287-293, 288-294, 291-297, 292-298, 295-301,
298-304, 305-311, 311-317, 315-321, 319-325, 322-328, 323-329,
325-331, 343-349, 354-360, 356-362, 374-380, 381-387, 383-389,
387-393, 389-395, 395-401

Cell attachment sequence.

amino acids 301-304

FIGURE 53

GGAGAAGAGGTTGTGTGGGACAAGCTGCTCCCGACAGAAGGATGTCGCTGCTGAGCCTGCCCTGG
CTGGGCCTCAGACCGGTGGCAATGTCCCATGGCTACTCCTGCTGCTGGTTGTGGGCTCCTGGCT
ACTCGCCCGCATCCTGGCTTGGACCTATGCCTTCTATAACAACACTGCCCGGCTCCAGTGTTC
CACAGCCCCAAAACGGAACGGTGGTGGGGTACCTGGGCCTGATCACTCCTACAGAGGAGGC
TTGAAGACTCGACCCAGATGTCGGCCACCTATTCCCAGGGCTTTACGGTATGGCTGGGTCCCAT
CATCCCCTTCATCGTTTTATGCCACCCTGACACCATCCGGTCTATCACCAATGCCTCAGCTGCCA
TTGCACCCAAGGATAATCTTTCATCAGGTTCTGAAGCCCTGGCTGGGAGAAGGGATACTGCTG
AGTGGCGGTGACAAGTGGAGCCGCCACCCTCGGATGCTGACGCCCGCCTTCCATTTCAACATCCT
GAAGTCTATATAACGATCTTCAACAAGAGTGCAAACATCATGCTTGACAAGTGGCAGCACCTGG
CCTCAGAGGGCAGCAGTCGCTGGACATGTTGAGCACATCAGCCTCATGACCTTGGACAGTCTA
CAGAAATGCATCTTCAGCTTTGACAGCCATTGTGAGGAGAGGCCAGTGAATATATGCCACCAT
CTTGGAGCTCAGTGCCCTTGTAGAGAAAAGAAGCCAGCATATCCTCCAGCACATGGACTTTCTGT
ATTACCTCTCCCATGACGGGCGGGCTTCCACAGGGCCTGCCGCCTGGTGCATGACTTCACAGAC
GCTGTTCATCCGGGAGCGGGTGCACCCCTCCCACTCAGGGTATGATGATTTTTTCAAAGACAA
AGCCAAGTCCAAGACTTTGGATTTTATTGATGTGCTTCTGCTGAGCAAGGATGAAGATGGGAAGG
CATTGTGAGATGAGGATATAAGAGCAGAGGCTGACACCTTCATGTTTGGAGGCCATGACACCAG
GCCAGTGGCCTCTCCTGGGTCTGTACAACCTTGGCAGGCACCCAGAATACCAGGAGCGCTGCCG
ACAGGAGGTGCAAGAGCTTCTGAAGGACCGCGATCCTAAAGAGATTGAATGGGACGACCTGGCCC
AGCTGCCCTTCTGACCATGTGCGTGAAGGAGAGCCTGAGGTTACATCCCCAGCTCCCTTCATC
TCCCAGTGTGCACCCAGGACATTGTTCTCCAGATGGCCGAGTCATCCCCAAGGCATTACCTG
CCTCATCGATATTATAGGGTCCATCACAACCAACTGTGTGGCCGGATCCTGAGGTCTACGACC
CCTTCCGCTTTGACCCAGAGAACAGCAAGGGGAGGTCACCTCTGGCTTTTATTCCTTTCTCCGCA
GGGCCAGGAACTGCATCGGGCAGGGTTCGCCATGGCGGAGATGAAAGTGGTCTGGCGTTGAT
GCTGCTGCACTTCCGGTTCCTGCCAGACCACACTGAGCCCCGAGGAAGCTGGAATTGATCATGC
GCGCCGAGGGCGGGCTTTGGCTGCGGGTGGAGCCCTGAATGTAGGCTTGCACTGACTTTCTGAC
CCATCCACCTGTTTTTTTGCAGATTGTCATGAATAAAACGGTGTCTGCTCAA

FIGURE 54

MSLLSLPWLGLRPVAMSPWLLLLLVVGSWLLARILAWTYAFYNNCRRLQCFPPKRNWFWGHLG
LITPTEEGLKDSTQMSATYSQGFTVWLGPIIPFIVLCHPDTIRSITNASAAIAPKDNLFIRFLKP
WLGEGILLSGGDKWSRHRMLTPAFHFNIKSYITIFNKSANIMLDKWQHLASEGSSRLDMFEHI
SLMTLDSLQKCIFSDSHCQERPSEYIATILELSALVEKRSQHILQHMDFLYLYLSDGRRFHRAC
RLVHDFTDAVIRERRRTLPTQGIDDFKDKAKSKTLDLFDVLLLSKDEDGKALSDEDIRAEADTF
MFGGHDTTASGLSWVLYNLAHPEYQERCQEVQELLKDRDPKEIEWDDLAQLPFLTMCVKESLR
LHPPAPFISRCCTQDIVLPDGRVIPKGITCLIDIIGVHHNPTVWPDPEVYDPFRFDPENSKGRSP
LAFIPFSAGPRNCIGQAFAMAEMKVVLALMLLHFRFLPDHTEPRRKLELIMRAEGGLWLRVEPLN
VGLQ

Important features:

Transmembrane domains:

amino acids 13-32 (type II), 77-102

Cytochrome P450 cysteine heme-iron ligand signature.

amino acids 461-471

N-glycosylation sites.

amino acids 112-116, 168-172

FIGURE 56

MGPVKQLKRMFEPTRLIATIMVLLCFALTLCSAFWWHNKGLALIFCILQSLALTWYLSFIPFAR
DAVKKCFVCLA

Important features:

Signal peptide:

amino acids 1-33

Type II fibronectin collagen-binding domain protein.

amino acids 30-72

FIGURE 57

CGGCTCGAGCTCGAGCCGAATCGGCTCGAGGGGCGAGTGGAGCACCAGCAGGCCGCCAACATGCTCTGTCTGTGCCTG
TACGTGCCGGTTCATCGGGGAAGCCAGACCGAGTCCAGTACTTTGAGTCGAAGGGGCTCCCTGCCGAGCTGAAGTCC
ATTTTCAAGCTCAGTGTCTTTCATCCCTCCAGGAATTCCTCCACTACCGCCAGTGGAAAGCAGAAAATTTGACAAGCT
GGAGATAAGGACCTTGATGGGCAGCTAGACTTTGAAGAATTTGFCATTATCTCCAAGATCATGAGAAGAAGCTGAGG
CTGGTGTAAAGATTTTGGACAAAAGAATGATGGACGCATTGACGCGCAGGAGATCATGCAGTCCCTGCCGGACTTG
GGAGTCAAGATATCTGAACAGCAGGCAGAAAAAATTCCTCAAGAGCATGGATAAAAACGGCAGATGACCATCGACTGG
AACGAGTGGAGAGACTACCACCTCCTCCACCCCGTGGAAAACATCCCCGAGATCATCTCTACTGGAAGCATTCCACG
ATCTTTGATGTGGGTGAGAATCTAACGGTCCCGGATGAGTTCACAGTGGAGGAGAGGCAGACGGGGATGTGGTGGAGA
CACCTGGTGGCAGGAGGTGGGGCAGGGCCGTATCCAGAACCTGCACGGCCCCCTGGACAGGCTCAAGTGCTCATG
CAGGTCCATGCCTCCCGCAGCAACAACATGGGCATCGTTGGTGGCTTCACTCAGATGATTGAGAAGGAGGGGCCAGG
TCACTTGGCGGGCAATGGCATCAAGCTCCTCAAAATGGCCCCGAATCAGCCATCAAAATTCATGGCTATGAGCAG
ATCAAGCCCTTGTGGTAGTGACCCAGGAGACTCTGAGGATTCACGAGAGGCTTGTGGCAGGGTCCCTGGCAGGGCC
ATCGCCAGAGCAGCATCTACCAATGGAGTCCCTGAAGACCCGGATGGCGATGGCGAAGACAGGCCAGTACTCAGGA
ATGCTGGACTGCGCCAGGAGGATCCTGGCCAGAGAGGGGGTGGCCGCTTCTACAAAGGCTATGTCCCAACATGCTG
GGCATCATCCCTATGCCGGCATCGACCTTGCAGTCTACGAGACGCTCAAGAATGCCTGGCTGCAGCATTATGCAGTG
AACAGCGCGGACCCCGCGTGTGTGCTCCTGGCCTGTGGCACCATGTCCAGTACCTGTGGCCAGCTGGCCAGCTAC
CCCCGGCCCTAGTCAGGACCCGGATGCAGGCGCAAGCCTCTATTGAGGGCGCTCCGGAGGTGACCATGAGCAGCCT
TTCAAACATATCCTGCGGACCGAGGGGGCCTTCGGGCTGTACAGGGGGCTGGCCCCAACTTCATGAAGGTCAATCCCA
GCTGTGAGGATCAGTACGTGGTCTACGAGAACCTGAAGATCACCTGGGCGTGCAGTCCGGGTGACGGGGGAGGGC
CGCCCGGAGTGGACTCGCTGATCCTGGGCGCAGCCTGGGGTGTGCAGCCATCTCATCTGTGAATGTGCCAACAT
AAGCTGTCTCGAGCCAGCTGTGAAAACCTTAGACGCACCCCGAGGAGGGTGGGGAGAGCTGGCAGGGCCAGGGCT
GTCCTGTGACCCAGCAGACCTCCTGTTGGTTCACGCAAGACACAGGCATTCCTTAGGGTCCAGGGTCCAGCAGG
CTCCGGGCTCACATGTGTAAGGACAGGACATTTCTGCAGTGCCTGCCAATAGTGAAGCTGGAGCCTGGAGCCGGCT
TAGTCTTCCATTTCACTTCGACCCAGCTGTGGCCACGGCCCCCTGCCCTCTGGTCTGCCGTGCATCTCCCTGTGC
CCTCTGTGCTGCCCTGTCTGCTGAGTAAGGTGGGAGGAGGGCTACAGCCACATCCCAACCCCTCGTCCAATCCC
ATAATCCATGATGAAAGGTGAGGTGAGTGGCCTCCAGGCCCTGACTTCCCAACCTACAGCATTGACGCCAATGCG
TGTGAAGGAAGAGAAAGGATCTGGCCTTGTGGTCACTGGCATCTGAGCCCTGCTGATGGCTGGGGCTCTCGGGCATG
CTTGGGATGTCAGGGGGCTCGGGCTGCCTGGCCTGGCTGCACAGAAGGCAAGTGTGGGGCTCATGGTGTCTGAGCT
GGCCTGGACCCTGTGAGGATGGGCCCCACCTCAGAACCAACTCACTGTCCCACTGTGGCATGAGGGCAGTGGAGCA
CCATGTTTGAAGGCGAAGGGCAGAGCCTTTGTGTCTTGGGGAGGGAAGGAAAAGGTGTTGGAGGCTTAATTTATGG
ACTGTTGGGAAAAGGGTTTTGTCAGAAAGGACAGCCGACAAATGAGCAGCTTCTGTGCTTCCAGAGGAGACGAGG
GAGCAGGAGCTTGGCTGACTGCTCAGAGTCTGTTCTGACGCCCTGGGGGTTCTGTCCAACCCAGCAGGGGGCCAGC
GGGACAGCCCCACATTCCTTGTGCTGCTTGGAACTATTTATTTGATTTATTTGAACAGAGTTATGCTCT
AATATTTTTATAGATTTGTTAATTAATAGCTTGTCAATTTCAAGTTCATTTTTTATTCATATTATGTTATGTT
GATTGTACCTTCCCAAGCCCGCCAGTGGGATGGGAGGAGGAGGAGAAGGGGGCCTTGGGCCGCTGCAGTCACTCT
GTCCAGAGAAATTCCTTTGGGACTGGAGGAGAAAAGCGCCAGAAGGCAGCAGCCCTGGCTCCTTTCTTTGGCAG
GTTGGGGAAGGGCTTGGCCCCAGCCTTAGGATTTGAGGTTTCACTGGGGGCTGGAGAGAGAGGGAGAACCTCAAT
AACCTTGAAGGTGGAATCCAGTTATTTCTGCGCTGCCAGGGTTTCTTTATTTCACTTTTTCTGAATGTCAAGGCAG
TGAGGTGCCTCTCACTGTGAATTTGTGGTGGGCGGGGCTGGAGGAGAGGGTGGGGGGCTGGCTCCCTCCCTCCAGC
CTTCTGCTGCCCTTGTAAACAATGCCGGCCAACTGGCGACCTCACGGTTGCATTCATTCCACAGAAATGACCTGA
TGAGGAAATCTTCAATAGGATGCAAAGATCAATGCAAAAATTTGTTATATATGAACATATAACTGGAGTCGTAAAAAG
CAAATTAAGAAAAGAAATGGACGTTAGAAGTTGTCAATTTAAAGCAGCCTTCTAATAAAGTTGTTTCAAAGCTGAAAAA
AA

FIGURE 58

MLCLCLYVPVIGEAEQTEFQYFESKGLPAELKSIKLSVFIQEFSTYRQWKQKIVQAGDKDLLDG
QLDFEEFVHYLQDHEKKLRLVFKILDKKNDGRIDAQEIFMQSLRDLGVKISEQQAEKILKSMDKNG
TMTIDWNEWRDYHLLHPVENIPEIILYWKHSTIFDVGENTVPDEFTEERQTMWRRHLVAGGG
AGAVSRTCTAPLDRLKVLQVHASRSNNMGIVGGFTQMIREGGARSLWRGNGINVLKIAPESAIAK
FMAYEQIKRLVGSQDETLEIHERLVAGSLAGAIQSSIYPMEVLKTRMALRKTGOYSGMLDCARR
ILAREGVAAFYKGYVPMMLGIIIPYAGIDLAVYETLKNWLOHYAVNSADPGVFVLLACGTMSSC
GQLASYPLALVTRMQAASIEGAPEVTMSSLFKHILRTEGAFGLYRGLAPNFMKVIPAVSISYV
VYENLKITLGVQSR

Important features:

Signal peptide:

amino acids 1-16

Putative transmembrane domains:

amino acids 284-304, 339-360, 376-394

Mitochondrial energy transfer proteins signature.

amino acids 206-215, 300-309

N-glycosylation sites.

amino acids 129-133, 169-173

Elongation Factor-hand calcium-binding protein.

amino acids 54-73, 85-104, 121-140

FIGURE 59

GGAAGGCAGCGGCAGCTCCACTCAGCCAGTACCCAGATACGCTGGGAACCTTCCCCAGCCATGGC
TTCCTGGGGCAGATCCTCTTCTGGAGCATAATTAGCATCATCATTATTCTGGCTGGAGCAATTG
CACTCATCATTTGGCTTTGGTATTTTCAGGGAGACACTCCATCACAGTCACTACTGTGCGCTCAGCT
GGGAACATTGGGGAGGATGGAATCCTGAGCTGCACCTTTGAACCTGACATCAAACCTTTCTGATAT
CGTGATACAATGGCTGAAGGAAGGTGTTTTAGGCTTGGTCCATGAGTTCAAAGAAGGCCAAAGATG
AGCTGTGCGGAGCAGGATGAAATGTTTCAGAGGCCGGACAGCAGTGTGCTGATCAAGTGATAGTT
GGCAATGCCTCTTTGCGGCTGAAAAACGTGCAACTCACAGATGCTGGCACCTACAAATGTTATAT
CATCACTTCTAAAGGCAAGGGGAATGCTAACCTTGAGTATAAACTGGAGCCTTCAGCATGCCGG
AAGTGAATGTGGACTATAATGCCAGCTCAGAGACCTTGCGGTGTGAGGCTCCCCGATGGTTCCCC
CAGCCCACAGTGGTCTGGGCATCCCAAGTTGACCAGGGAGCCAACCTTCTCGGAAGTCTCCAATAC
CAGCTTTGAGCTGAACTCTGAGAATGTGACCATGAAGGTTGTGTCTGTGCTCTACAATGTTACGA
TCAACAACACATACTCCTGTATGATTGAAAATGACATTGCCAAAGCAACAGGGGATATCAAAGTG
ACAGAATCGGAGATCAAAGGGCGGAGTCACTACAGCTGCTAAACTCAAAGGCTTCTCTGTGTGT
CTCTTCTTTCTTTGCCATCAGCTGGGCACTTCTGCCTCTCAGCCCTTACCTGATGCTAAAATAAT
GTGCCTTGCCACAAAAAGCATGCAAAGTCATTGTTACAACAGGGATCTACAGAACTATTTTAC
CACCAGATATGACCTAGTTTTATATTTCTGGGAGGAAATGAATTCATATCTAGAAGTCTGGAGTG
AGCAAACAAGAGCAAGAAACAAAAAGAAGCCAAAAGCAGAAGGCTCCAATATGAACAAGATAAAT
CTATCTTCAAAGACATATTAGAAGTTGGGAAAATAATTCATGTGAAGTACAGCAAGTGTGTTAAGA
GTGATAAGTAAATGCACGTGGAGACAAGTGCATCCCAGATCTCAGGGACCTCCCCCTGCCTGT
CACCTGGGGAGTGAGAGGACAGGATAGTGCATGTTCTTTGTCTCTGAATTTTTAGTTATATGTGC
TGTAATGTTGCTCTGAGGAAGCCCCCTGGAAAGTCTATCCCAACATATCCACATCTTATATCCAC
AAATTAAGCTGTAGTATGTACCCTAAGACGCTGCTAATTGACTGCCACTTCGCAACTCAGGGGCG
GCTGCATTTTAGTAATGGGTCAAATGATTCACTTTTATGATGCTTCAAAGGTGCCTTGGCTTC
TCTTCCCAACTGACAAATGCCAAAGTTGAGAAAAATGATCATAATTTTAGCATAAACAGAGCAGT
CGGGGACACCGATTTTATAAATAAACTGAGCACCTTCTTTTAAACAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 60

MASLGQILFWSIISIIIIILAGAIALIIGFGISGRHSITVTTVASAGNIGEDGILSCTFEPDIKLS
DIVIQWLKEGVLGLVHEFKEGKDELSEQDEMFRGRTAVFADQVIVGNASLRLKNVQLTDAGTYKC
YIITSKGKGNANLEYKTGAFSMPEVNVVDYNASSETLRCEAPRWFPPQPTVVWASQVDQGANFSEVS
NTSFELNSENVTMKVVSVLYNVTINNTYSCMIENDIAKATGDIKVTESEIKRRSHLQLLNSKASL
CVSSFFAISWALLPLSPYMLK

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 258-281

N-glycosylation sites.

amino acids 112-116, 160-164, 190-194, 196-200, 205-209, 216-220,
220-224

N-myristoylation sites.

amino acids 52-58, 126-132, 188-194

FIGURE 61

TGACGTCAGAATCACCATGGCCAGCTATCCTTACCGGCAGGGCTGCCAGGAGCTGCAGGACAAG
CACCAGGAGCCCCTCCGGGTAGCTACTACCTGGACCCCCAATAGTGGAGGGCAGTATGGTAGT
GGGCTACCCCCTGGTGGTGGTTATGGGGTCTGCCCTGGAGGGCCTTATGGACCACCAGCTGG
TGAGGGGCCCTATGGACACCCCAATCCTGGGATGTTCCCTCTGGAActCCAGGAGGACCATATG
GCGGTGCAGCTCCC GGGGCCCCCTATGGTCAGCCACCTCCAAGTTCTACGGTGCCAGCAGCCT
GGGCTTTATGGACAGGGTGGCGCCCCTCCAATGTGGATCCTGAGGCCTACTCTGGTTCCAGTC
GGTGGACTCAGATCACAGTGGCTATATCTCCATGAAGGAGCTAAAGCAGGCCCTGGTCAACTGCA
ATTGGTCTTCATTCAATGATGAGACCTGCCTCATGATGATAAACATGTTTGACAAGACCAAGTCA
GGCCGCATCGATGTCTACGGCTTCTCAGCCCTGTGAAATTCATCCAGCAGTGAAGAACCCTCTT
CCAGCAGTATGACCGGGACCGCTCGGGCTCCATTAGCTACACAGAGCTGCAGCAAGCTCTGTCCC
AAATGGGCTACAACCTGAGCCCCAGTTCACCCAGCTTCTGGTCTCCCGCTACTGCCACGCTCT
GCCAATCCTGCCATGCAGCTTGACCGCTTCATCCAGGTGTGCACCCAGCTGCAGGTGCTGACAGA
GGCCTTCCGGGAGAAGGACACAGCTGTACAAGGCAACATCCGGCTCAGCTTCGAGGACTTCGTCA
CCATGACAGCTTCTCGGATGCTATCACCCAACCATCTGTGGAGAGTGGAGTGCACCAGGGACCTT
TCCTGGCTTCTTAGAGTGAGAGAAGTATGTGGACATCTCTTCTTTTCTGTCCCTCTAGAAGAAC
ATTCTCCCTTGCTTGATGCAACACTGTTCAAAAGAGGGTGGAGAGTCTGCATCATAGCCACCA
AATAGTGAGGACCGGGCTGAGGCCACACAGATAGGGCCCTGATGGAGGAGAGGATAGAAGTTGA
ATGTCCTGATGGCCATGAGCAGTTGAGTGGCACAGCCTGGCACCAGGAGCAGGTCTTGTAAATGG
AGTTAGTGTCCAGTCAGCTGAGCTCCACCCTGATGCCAGTGGTGGAGTGTTCATCGGCCTGTACC
GTTAGTACCTGTGTTCCCTCACCAGGCCATCCTGTCAAACGAGCCATTTTCTCCAAAGTGAAT
CTGACCAAGCATGAGAGAGATCTGTCTATGGGACCAGTGGCTTGGATTCTGCCACCCATAAAT
CCTTGTGTGTTAACTTCTAGCTGCCGTTGGGCTGGCCCTGCTCAGACAAATCTGCTCCCTGGGCAT
CTTTGGCCAGGCTTCTGCCCCCTGCAGCTGGGACCCCTCACTTGCTGCCATGCTCTGCTCGGCT
TCAGTCTCCAGGAGACAGTGGTCACCTCTCCCTGCCAATACTTTTTTAATTTGCATTTTTTTTC
ATTTGGGGCCAAAAGTCCAGTGAATTGTAAGCTTCAATAAAAGGATGAAACTCTGA

FIGURE 62

MASYPYRQGCPGAAGQAPGAPPGSYYPGPPNSGGQYGSGLPPGGGYGGPAPGGPYGPPAGGGPYG
HPNPGMFPSGTPGGPYGGAAPGGPYGQPPSSYGAQQPGLYGQGGAPPNVDPEAYSWFQSVSDSH
SGYISMKELKQALVNCNWSSFNDETCLMMINMFDKTKSGRIDVYGFSALWKFIQQWKNLFQQYDR
DRSGSISYTELQQALSQMGYNLSPQFTQLLVSRYCPRSANPAMQLDRFIQVCTQLQVLTEAFREK
DTAVQGNIRLSFEDFVMTASRML

Important features of the protein:

Signal peptide:

amino acids 1-19

N-glycosylation site.

amino acids 147-150

Casein kinase II phosphorylation sites.

amino acids 135-138, 150-153, 202-205, 271-274

N-myristoylation sites.

amino acids 9-14, 15-20, 19-24, 33-38, 34-39, 39-44, 43-48, 61-
66, 70-75, 78-83, 83-88, 87-92, 110-115

FIGURE 63

CAGGATGCAGGGCCGCGTGGCAGGGAGCTGCGCTCCTCTGGGCCTGCTCCTGGTCTGTCTTCATC
TCCCAGGCCTCTTTGCCCGGAGCATCGGTGTTGTGGAGGAGAAAGTTTCCCAAACCTTCGGGACC
AACTTGCCTCAGCTCGGACAACCTTCCCTCCACTGGCCCTCTAACTCTGAACATCCGCAGCCCGC
TCTGGACCCTAGGTCTAATGACTTGGCAAGGGTTCCCTCTGAAGCTCAGCGTGCCTCCATCAGATG
GCTTCCCACCTGCAGGAGGTTCTGCAGTGCAGAGGTGGCCTCCATCGTGGGGGCTGCCTGCCATG
GATTCTGGCCCCCTGAGGATCCTTGGCAGATGATGGCTGCTGCGGCTGAGGACCGCTGGGGGA
AGCGCTGCCTGAAGAATCTCTTACCTCTCCAGTGTGCGGCCCTCGCTCCGGGCAGTGGCCCTT
TGCCTGGGGAGTCTTCTCCCGATGCCACAGGCCTCTCACCTGAGGCTTCACTCCTCCACCAGGAC
TCGGAGTCCAGACGACTGCCCGTTCTAATTCAGTGGGAGCCGGGGGAAAAATCCTTTCCCAACG
CCCTCCCTGGTCTCTCATCCACAGGGTTCTGCCTGATCACCCCTGGGGTACCCTGAATCCCAGTG
TGTCTGGGGAGGTGGAGGCCCTGGGACTGGTTGGGGAACGAGGCCCATGCCACACCCTGAGGGA
ATCTGGGGTATCAATAATCAACCCCCAGGTACCAGCTGGGGAAATATTAATCGGTATCCAGGAGG
CAGCTGGGGAAATATTAATCGGTATCCAGGAGGCAGCTGGGGAAATATTAATCGGTATCCAGGAG
GCAGCTGGGGAAATATTCATCTATACCCAGGTATCAATAACCCATTTCTCCTGGAGTTCTCCGC
CCTCCTGGCTCTTCTTGGAACATCCCAGCTGGCTTCCCTAATCCTCCAAGCCCTAGGTTGCAGTG
GGGCTAGAGCACGATAGAGGAAACCCAAACATTGGGAGTTAGAGTCTGCTCCCGCCCTTGCTG
TGTGGGCTCAATCCAGGCCCTGTTAACATGTTTCCAGCACTATCCCCACTTTTCAGTGCCTCCCC
TGCTCATCTCCAATAAAAATAAAAGCACTTATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AA

FIGURE 64

MQGRVAGSCAPLGLLLVCLHLPGLFARSIGVVEEKVSQNFGTNLPQLGQPSSTGPSNSEHPQPAL
DPRSNDLARVPLKLSVPPSDGFPPAGGSAVQRWPPSWGLPAMDSWPPEDPWQMAAAAEDRLGEA
LPEELSYLSSAAALAPGSGPLPGESSPDATGLSPEASLLHQDSESRRLPRNSNLGAGGKILSORP
PWSLIHRVLPDHPWGTLNPSVSWG GGGPGTGWGTRMPHPEGIWGINNQPPGTSWGNINRYPGGS
WGNINRYPGGSWGNINRYPGGSWGNIHLYPGINNPFPVLRPPGSSWNI PAGFPNPPSPRLQWG

Important features of the protein:

Signal peptide:

amino acids 1-26

Casein kinase II phosphorylation sites.

amino acids 56-59, 155-158

N-myristoylation sites.

amino acids 48-53, 220-225, 221-226, 224-229, 247-252, 258-263,
259-264, 269-274, 270-275, 280-285, 281-286, 305-310

FIGURE 66

MGSG LPLV LLL TLL GSS HGT GPG MTL QLKL KES FLT NSS YESS FLE LLE KLCL LLL HLP SGTS VTL
HHARSQH HVCNT

Important features:

Signal peptide:

amino acids 1-19

N-glycosylation site.

amino acids 37-41

N-myristoylation sites.

amino acids 15-21, 19-25, 60-66

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

FIGURE 67

ACGGACCGAGGGTTCGAGGGAGGGACACGGACCAGGAACCTGAGCTAGGTCAAAGACGCCCGGGC
CAGGTGCCCCGTCGCAGGTGCCCTGGCCGAGATGCGGTAGGAGGGGCGAGCGCGAGAAGCCCC
TTCCTCGGGCGTCGAACCCGCCACCCAGCCCAATGGCGAACCCCGGGCTGGGGCTGCTTCTGGCG
CTGGGCCTGCCGTTCTCTGCTGGCCCGCTGGGGCCGAGCCTGGGGGCAAATACAGACCACTTCTGC
AAATGAGAATAGCACTGTTTTGCCTTCATCCACCAGCTCCAGCTCCGATGGCAACCTGCGTCCGG
AAGCCATCACTGCTATCATCGTGGTCTTCTCCCTCTTGGCTGCCTTGCTCCTGGCTGTGGGGCTG
GCACTGTTGGTGC GGAAGCTTCGGGAGAAGCGGCAGACGGAGGGCACCTACCGGCCAGTAGCGA
GGAGCAGTTCTCCCATGCAGCCGAGGCCCGGGCCCTCAGGACTCCAAGGAGACGGTGCAGGGCT
GCCTGCCCATCTAGGTCCCCTCTCCTGCATCTGTCTCCCTTCATTGCTGTGTGACCTTGGGGAAA
GGCAGTGCCCTCTCTGGGCAGTCAGATCCACCCAGTGCTTAATAGCAGGGAAGAAGGTACTCAA
AGACTCTGCCCTGAGGTCAAGAGAGGATGGGGCTATTCACTTTTATATATTTATATAAAATTAG
TAGTGAGATGTAAAAAAAAAAAAAAAAAAAA

FIGURE 68

MANPGLGLLLALGLPFLLRWGRAWGQIQTTSANENSTVLPSSSTSSSSDGNLRPEAITAIIVVFS
LLAALLLAVGLALLVRKLRKREKQTEGTYRPSSEEQFSHAAEARAPQDSKETVQGCLPI

Important features:

Signal peptide:

amino acids 1-19

Transmembrane domain:

amino acids 56-80

N-glycosylation site.

amino acids 36-40

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 86-90

Tyrosine kinase phosphorylation site.

amino acids 86-94

N-myristoylation sites.

amino acids 7-13, 26-32

FIGURE 69

GCCAGGAATAACTAGAGAGGAACAATGGGGTTATTTCAGAGGTTTTGTTTTCTCTAGTTCTGTGCCTGCTGCCACCAG
TCAAATACTTCTTCAATTAAGCTGAATAATAATGGCTTTGAAGATATTGTCAATTGTTATAGATCCTAGTGTGCCAGAA
GATGAAAAATAATTAACAAATAGAGGATATGGTACTACAGCTTCTACGTACCTGTTTGAAGCCACAGAAAAAGA
TTTTTTTTCAAAAATGTATCTATATTAATTCCTGAGAATTGGAAGGAAAAATCCTCAGTACAAAAGGCCAAAACATGAA
AACATAAACATGCTGATGTTATAGTTGCACCACCTACACTCCAGGTAGAGATGAACCATAACCAAGCAGTTCCACA
GAATGTGGAGAGAAAGGCGAATACATTCACTTCAACCCCTGACCTTCTACTTGGAAAAAACAAAATGAATATGGACCA
CCAGGCCAACTGTTTGTCCATGAGTGGGCTCACCTCCGGTGGGAGTGTGATGAGTACAATGAAGATCAGCCTTTC
TACCGTGCTAAGTCAAAAAAATCGAAGCAACAAGGTGTTCCGAGGTATCTCTGGTAGAAATAGAGTTTATAAGTGT
CAAGGAGGCGCTGTCTTAGTAGAGCATGCAGAAATGATTTCTACAACAAAATGATGGAAAAGATTGTCAATTTCTTT
CCTGATAAAGTACAACAGAAAAAGCATCCATAATGTTTATGCAAAGTATTGATTTCTGTTGTAATTTTGTAAACGAA
AAAACCCATAATCAAGAAGCTCCAAGCCTACAAAACATAAAGTGCAATTTTAGAAGTACATGGGAGGTGATTAGCAAT
TCTGAGGATTTAAAAACCCATACCCATGGTGACACCACCTCCTCCACCTGCTTCTCATTGCTGAAGATCAGTCAA
AGAATTGTGTGCTTAGTCTTGATAAGTCTGGAAGCATGGGGGTAAGGACCCGCTAAATCGAATGAATCAAGCAGCA
AAACATTTCTGCTGCAGACTGTTGAAAATGGATCCTGGTGGGGATGGTTCACTTTGATAGTACTGCCACTATTGTA
AATAAGCTAATCAAATAAAAAGCAGTATGAAAGAAACACACTCATGGCAGGATTACCTACATATCCTCTGGGAGGA
ACTTCCATCTGCTCTGGAATTAATATGCATTTCAAGGTGATTGGAGAGCTACATTTCCCAACTCGATGGATCCGAGTA
CTGCTGCTGACTGATGGGAGGATAACACTGCAAGTCTTGTATTGATGAAGTAAAACAAAGTGGGGCCATTGTTCAAT
TTTTATTGCTTTGGGAAGAGCTGCTGATGAAGCAGTAATAGAGATGAGCAAGATAACAGGAGGAAGTCAATTTTATGTT
TCAGATGAAGCTCAGAACATGGCCTCATTGATGCTTTTGGGGCTCTTACATCAGGAAATACTGATCTCTCCAGAAG
TCCCTTCAGCTCGAAAAGTAAGGGATTAAACACTGAATAGTAATGCCTGGATGAACGACACTGTCATAAATGATAGTACA
GTGGGAAAGGACACGTTCTTCTCATCACATGGAACAGTCTGCCCTCCAGTATTCTCTCTGGGATCCAGTGGGAACA
ATAATGAAAAATTTACAGTGGATGCAACTTCCAAAATGGCCTATCTCAGTATTCCAGGAACTGCAAAGTGGGCACT
TGGGCATACAATCTTCAAGCCAAAGCGAACCCAGAACATTAACATTTACAGTAACCTTCTCGAGCAGCAAAATCTTCT
GTGCCTCCAATCACAGTGAATGCTAAAATGAATAAGGACGTAACAGTTCCTCCAGCCCAATGATTTGTTTACGCAGAA
ATTTACAAAGGATATGACCTGTTCTTGGAGCCAATGTGACTGCTTTTCAATGAAATCACAGAATGGACATACAGAAGTT
TTGGAACTTTGGATAATGTTGCAAGGCGCTGATTCTTCAAGAATGATGGAGTCTACTCCAGGTATTTTACAGCATAT
ACAGAAAATGGCAGATATAGCTTAAAAGTTCGGGCTCATGGAGGAGCAAACACTGCCAGGCTAAAATTACGGCCTCCA
CTGAATAGAGCCCGGTACATACCAGGCTGGGTAGTGAACGGGGAAATGAAAGCAAACCCGCAAGACCTGAAATTTGAT
GAGGATCTCAGACCCTTGGAGGATTTCAAGCCGACAGCATCCGAGGTGCATTTGTTGATCACAAGTCCCAAGC
CTTCCCTTGCCCTGACCAATACCCACCAAGTCAAATCACAGACCTTGTATGCCACAGTTTATGAGGATAAGATTTATCTT
ACATGGACAGCACCAAGGAGATAATTTTATGATGTTGGAAAAGTCAACGTTATATCATAAGAATAAGTGAAGTATTCTT
GATCTAAGAGACAGTTTGTATGATGCTCTTCAAGTAAATACTACTGATCTGTCACCAAAGGAGGCCAACTCCAAGGAA
AGCTTTGCATTTAAACCAGAAAATATCTCAGAAGAAAATGCAACCCACATATTTATTGCCATTAAGATATAGATAAA
AGCAATTTGACATCAAAGTATCCAACATGCACAAGTAACTTTGTTTATCCCTCAAGCAATCCTGATGACATTGAT
CCTACACCTACTCCTACTCCTACTCCTACTCTGATAAAAAGTCAATAATCTGGAGTTAATATTTCTACGCTGGTATTG
TCTGTGATGGGTCTGTTGTAATTTGTTAACTTTATTTTAAAGTACCACCATTGAACTTAAACGAAAGAAAAAATCTTC
AAGTAGACCTAGAAGAGAGTTTTAAAAAACAAAACATGTAAGTAAAGGATATTTCTGAATCTTAAAAATTCATCCCAT
GTGTGATCATAAACTCATAAAAATAATTTAAGATGTCGAAAAGGATACTTTGATTAAATAAAAACACTCATGGATA
TGTAAAAATGTCAGATTAATAATTAATAGTTTCATTTATTTGTTATTTTATTGTAAGAAATAGTGAACAAG
ATCCTTTTTCTACTGATACCTGGTTGTATTTATTTGATGCAACAGTTTTCTGAAATGATATTTCAAATGTCATCAA
GAAATAAAATCATCTATCTGAGTAGTCAAAATACAAGTAAAGGAGAGCAAATAAACACACTTTGGAAAAAATAAAAA
AA

10063728

FIGURE 70

MGLFRGFVFLVLCLLHQSNSTFIKLNNGFEDIVIVIDPSVPEDEKIEQIEDMVTASTYLFE
ATEKRFFFKNVSI LIPENWKENPOYKRPKHENHKKHADVIVAPPTLPGRDEPYTKQFTECGEKGEY
IHFTPDLLLGGKQNEYGPPGKLFVHEWAHLRWGVFDEYNEDQPFYRAKSKKIEATRC SAGISGRN
RVYKCGGSCLSRACRIDSTTKLYGKDCQFFPKVQTEKASIMFMQSIDSVVEFCNEKTHNQEAP
SLQNIKCNFRSTWEVISNSEDFKNTIPMVTPPPPPVFSLKISQRIVCLVLDKSGSMGGKDR LNR
MNQAAKHFL LQTVENG SWVGMVHF DSTATIVNKLIQIKSSDERNTLMAGLP TYPLGGTSICSGIK
YAFQVIGELHSQLDGSEVLL LLDGEDNTASSCIDEVKQSGAIVHFIALGRAADEAVIEMSKITGG
SHFYVSDEAQNNGLIDAFGALTSGNTDLSQKSLQLESKGLTLNSNAWMNDTVIIDSTV GKDTFFL
ITWNSLPPSISLWDPSGTIMENFTVDATSKMAYLSIPGTAKVGTWAYNLQAKANPETLTITVTSR
AANSSVPPITVNAKMNKDVNSFPPMIVYAEILQGYVPVLGANVTAFIESQNGHTEVLELLDNGA
GADSFKNDGVYSRYFTAYTENGRYSLKVAHGGANTARLKL RPPLNRAAYIPGWVNGEIEANPP
RPEIDEDTQT TLEDFSR TASSGGAFFVSVQVPSLPLPDQYPPS QITDL DATVHEDKIILTWTAPGDN
FDVGKVQRYIIRISASILD L RDSFDDALQVNTTDLSPKEANSKESFAFKPENISEENATHIFIAI
KSIDKSNLTSKVS NIAQVTLFIPQANPDDIDPTPTPTPTPTPKSHNSGVNISTLVLSVIGSVVI
VNFILSTTI

Signal peptide:

amino acids 1-21

Putative transmembrane domains:

amino acids 284-300, 617-633

Leucine zipper pattern.

amino acids 469-491, 476-498

N-glycosylation site.

amino acids 20-24, 75-79, 340-344, 504-508, 542-546, 588-592,
628-632, 811-815, 832-836, 837-841, 852-856, 896-900

FIGURE 71

CTCCTTAGGTGAAAACCCCTGGGAGTAGAGTACTGACAGCAAAGACCGGAAAGACCATACTGCCCGGGCAGGGGTGA
CAACAGGTGTATCTTTTGTATCTCGTGTGTGGCTGCCTTCCCTATTCAAGGAAAGACGCCAAGGTAATTTTGACCCA
GAGGAGCAATGATGTAGCCACCTCCTAACCTTCCCTTCTTGAACCCCAAGTTATGCCAGGATTTACTAGAGAGTGTCA
ACTCAACCAGCAAGCGGCTCCTTCGGCTTAACCTTGTGTTGGAGGAGAGAACCCTTGTGGGGCTGCCTTCTTAGCA
GTGCTCAGAAGTGAAGTGCCTGAGGGTGGACCAGAAGAAAGGAAAGGTCCTTCTGCTGTTGGCTGCACATCAGGAA
GGCTGTGATGGGAATGAAGGTGAAAACCTTGGAGATTTCACTTCAGTCATTGCTTCTGCCTGCAAGATCATCCTTTAAA
AGTAGAGAAGCTGCTCTGTTGGTGGTTAACTCCAAGAGGCAGAACTCGTTCTAGAAGGAAATGGATGCAAGCAGCTC
CGGGGGCCCCAAACGCATGCTTCTGTGGTCTAGCCAGGGAAGCCCTTCCGTGGGGGGCCCCGGCTTTGAGGGATGCC
ACCGGTTCTGGACGCATGGCTGATTCCTGAATGATGATGGTTCGCGGGGGCTGCTTGGCTGGATTTCCGGGGTGGT
GTTTTGCTGGTGCCTCTGCTGTGCTATCTCTGTCTGTACATGTTGGCTGCACCCCAAAGGTGACGAGGAGCAG
CTGGCACTCCCCAGGGCCAAACAGCCCCACGGGGAAGGAGGGTACCAGGCCCTCTCAGGAGTGGGAGGAGCAGCAC
CGCACTACGTGAGCAGCCTGAAGCGGCAGATCGCACAGCTCAAGGAGGAGCTGCAGGAGAGGAGTGAAGCAGCTCAGG
AATGGGCAGTACCAGCCAGCGATGCTGCTGGCCTGGTCTGGACAGGAGCCCCCAAGAGAAACCCAGGCCGACCTC
CTGGCTTCTGCACTCGCAGGTGGACAAGGCAGAGGTGAATGCTGGCCTCAAGCTGGCCACAGAGTATGCAGCAGTG
CCTTTTCGATAGCTTTACTCTACAGAAGGTGTACCAGCTGGAGACTGGCCTTACCCGCCACCCGAGGAGAGCCTGTG
AGGAAGGACAAGCGGGATGATTTGGTGAAGCCATTGAATCAGCCTTGGAGACCCTGAACAATCCTCGAGAGAACAGC
CCCAATCACCGTCTTACACGGCTCTGATTTTCATAGAAGGGATCTACCGAACAGAAAGGGACAAAGGGACATTTGTAT
GAGCTCACCTTCAAAGGGGACCAAAACACGAATTCAAACGGCTCATCTTATTTTCGACCAATTCAGCCCATCATGAAA
GTGAAAAATGAAAGCTCAACATGGCCAAACCGTTATCAATGTTATCGTGCCTTAGCAAAAAGGGTGGACAAGTTC
CGGCAGTTCATGCAGAAATTCAGGGAGATGTGCATTGAGCAGGATGGGAGAGTCCATCTCAGTGTGTTTACTTTGGG
AAAGAAGAAATAAATGAAGTCAAAGGAATACTTGAACAACCTTCAAAGCTGCCAACTTCAGGAACCTTACCTTCATC
CAGCTGAATGGAGAATTTTCTCGGGGAAAGGGACTTGTATGTTGGAGCCCGCTTCTGGAAGGGAAAGCAACCTCTCTC
TTTTTCTGTGATGTGGACATCTACTTCCACATCTGAATTCCTCAATACGTGTAGGCTGAATACACAGCCAGGGAAGAAG
GTATTTTATCCAGTCTTTTTCAGTCAGTACAATCTGGCATAATATACGGCCCATGATGCAGTCCCTCCCTTGGAA
CAGCAGCTGGTATAAAGAAGGAACTGGATTTTGGAGAGACTTTGGATTTGGGATGACGTGTGATATCGGTGAGCAG
TTCATCAATATAGGTGGGTTTGTCTGGACATCAAAGGCTGGGGCGGAGAGGATGTGCACCTTATCGCAAGTATCTC
CACAGCAACCTCATAGTGGTACGGACCGCTGTGCGAGGACTCTTCCACCTTGGCATGAGAAGCGCTGCATGGACGAG
CTGACCCCGGAGCAGTACAAGATGTGCATGCAGTCCAAGGCCATGAACGAGGCATCCACGGCCAGCTGGGCATGCTG
GTGTTCCAGGCACGAGATAGAGGCTCACCTTCGCAACAGAAACAGAAGACAAGTAGCAAAAAACATGAACTCCAGAG
GAAGGATTTGGGAGACACTTTTTCTTTCTTTTTCGAAATTAAGTGGCTGCAACAGAGAAAAGACTTCCATAAA
GGACGACAAAAGAAATGGACTGATGGGTGAGAGATGAGAAAGCCCTCGATTTCTCTCTGTTGGGCTTTTTTACAACAGA
AATCAAAATCTCCGCTTGGCTGCAAAAGTAACCCAGTTGCACCCCTGGAAGTGTCTGACAAAGGCAGAATGCTTGTG
AGATTATAAGCCTAATGGTGTGGAGGTTTGTATGGTGTACAAATACACTGAGACCTGTGTTTTGTTGCTCATTGA
AATATTATGATTTAAGAGCAGTTTTGTAATAAATTCATTAGCATGAAAGGCAAGCATATTTCTCCTCATATGAATGA
GCCTATCAGCAGGGCTCTAGTTTCTAGGAATGCTAAAATATCAGAAGGCAGGAGAGGATAGGCTTATTATGATACT
AGTGAGTACATTAAGTAAATAAATGACCAGAAAAGAAAAGAAACCATAAATATCGTGTATATTTTCCCCAAGAT
TAACCAAAATAAATCTGCTTATCTTTTTGGTGTCTTTTAACTGTCTCCGTTTTTTCTTTTATTTAAAATGCACT
TTTTTCCCTTGTGAGTTATAGTCTGCTTATTTAATACCACCTTGGCAAGCCTTACAAGAGAGCACAAGTTGGCCTAC
ATTTTTATATTTTTAAGAAGATACTTTGAGATGCATTATGAGAATTTTCAAGTCAAAGCATCAAATGATGCCATAT
CCAAGGACATGCCAAATGCTGATTCGTGAGGCACTGAATGTGAGGCAATTGAGACATAGGGAAGGAATGGTTGTACT
AATACAGACGTACAGATACTTCTCTGAAGAGTATTTTCAAGAGGAGCAACTGAACACTGGAGGAAAAGAAAATGAC
ACTTTCTGCTTTACAGAAAAGGAACTCATTACAGACTGGTATATCGTATGTACCTAAAAGTCAAGAAACCAATTTT
CTCCTCAGAAGTAGGACCGCTTCTTACCTGTTTAAATAAACCAAGTATACCGTGTGAACCAACAATCTCTTTT
AAAACAGGGTGTCTCCTGGCTTCTGGCTTCCATAAGAAGAAATGGAGAAAATATATATATATATATATATATTGT
GAAAATCAATCCATCTGCCAGAATCTAGTGGATGGAAGTTTTTGTCTACATGTTATCCACCCAGGCCAGGTGGAAG
TAACCTGAATATTTTTTAAATTAAGCAGTCTACTCAATCACAAGATGCTTCTGAAAATGCAATTTATACCATT
CAAATATTTTTTAAAATAAATACAGTTAACATAGAGTGGTTCCTTCAATTCATGTGAAAATTTATAGCCAGCACCAG
ATGCATGAGCTAATATCTCTTTGAGTCTTGTCTGCTTCTGCTCACAGTAAACTCATTGTTAAAAGCTTCAAGAAC
ATTCAAGCTGTGGTGTGTTAAAATAATGCATTGTATTGATTTGACTGGTAGTTTATGAAAATTTAATAAAACACAGG
CCATGAATGGAAGGTGGTATTGCACAGCTAATAAATAATGATTTGTGGATATGAA

FIGURE 72

MMVRRGLLAWISRVVLLVLLCCAISVLYMLACTPKGDEEQLALPRANSPTGKEGYQAVLQEW
EQHRNYVSSSLKRQIAQLKEELQERSEQLRNGQYQASDAAGLGLDRSPPEKTQADLLAFLHSQVDK
AEVNAGVKLATEYAAVPFDSFTLQKVYQLETGLTRHPEEKPVKDKRDELVEAIESALETLNPA
ENSPNHRPYTASDFIEGIYRTERDKGTYELTFKGDHKHEFKRLILFRPFSPIMVKNEKLNMAN
TLINVIVPLAKRVDFRQFMQNFREMCIEQDGRVHLTVVYFGKEEINEVKGILENTSKAANFRNF
TFIQLNGEFSRGKGLDVGARFWKGSNVLLFFCDVDIYFTSEFLNTCRLNTQPGKKVFPVLFQY
NPGIYGHDAVPPLEQQLVIKKETGFWRDFGFGMTCQYRSDFINIGGFDDIKWGGEDVHLYR
KYLHNLIVVRTPVRGLFHLWHEKRCMDELTPQYKCMQSKAMNEASHGQLGMLVFRHEIEAHL
RKQKQKTSKKT

Important features:

Signal peptide:

amino acids 1-27

N-glycosylation sites.

amino acids 315-319, 324-328

N-myristoylation sites.

amino acids 96-102, 136-142, 212-218, 311-317, 339-345, 393-399

Amidation site.

amino acids 377-381

FIGURE 74

MLFSALLLEVIWILAADGGQHWTYEGPHGQDHWPAASYPECGNNAQSPIDIQTDSVTFDPDLPALQ
PHGYDQPGTEPLDLHNNNGHTVQLSLPSTLYLGGGLPRKYVAAQLHLHWGQKGSPPGGSEHQINSEAT
FAELHIVHYDSDSYDSLSEAAERPQGLAVLGILIEVGETKNIAYEHILSHLHEVRHKDQKTSVPP
FNLRELLPKQLGQYFRYNGSLTTPPCYQSVLWTVFYRRSQISMEQLEKLGTLFSTEEEPSKLLV
QNYRALQPLNQRMVFASFIQAGSSYTTGEMLSLGVGILVGCCLLLAVYFIARKIRKKRLENRKS
VVFTSAQATTEA

Important features of the protein:

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 291-310

N-glycosylation site.

amino acids 213-216

Eukaryotic-type carbonic anhydrases proteins

amino acids 197-245, 104-140, 22-69

FIGURE 76

MSYNGLHQRVFKELKLLTLCSSISQIGPPEVALTTDEKSI SVVLTAPEKWKRNPEDLPVSMQQIY
SNLKYNVSVLNTKSNRTWSQCVTNHTLVLTWLEPNTLYCVHVESFVPGPPRAQPSEKQCARTLK
DQSSEFKAKIIFWYVLPISITVFLFSVMGYSIYRYIHVGKEKHPANLILYIGNEFDKRFVPAEK
IVINFITLNISSDDSKISHQDMSLLGKSSDVSSLNDPQPSGNLRPPQEEEEVKHLGYASHLMEIFC
DSEENTEGTSLTQQESLRTIPDKTVIEYDYVRTTDICAGPEEQELSLQEEVSTQGTLLLESQA
ALAVLGPQTLQYSYTPQLQDLPLAQEHTDSEEGPEEPPSTTLVDWDPQTGRLCIPSLSSFDQDS
EGCEPSEGDGLGEEGLLSRLYEPPAPDRPPGENETYLMQFMEEWGLYVQMEN

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 140-163

N-glycosylation sites.

amino acids 71-74, 80-83, 89-92, 204-207, 423-426

FIGURE 77

GAGGAGCGGGCCGAGGACTCCAGCGTGCCAGGTCTGGCATCCTGCACTTGCTGCCCTCTGACAC
CTGGGAAGATGCGCCGGCCCGTGGACCTTCACCCCTTCTCTGTGGTTTGCTGGCAGCCACCTTGATC
CAAGCCACCCTCAGTCCCACCTGCAGTTCTCATCCTCGGCCAAAAGTCATCAAAGAAAAGCTGAC
ACAGGAGCTGAAGGACCACAACGCCACCAGCATCCTGCAGCAGCTGCCGCTGCTCAGTGCCATGC
GGGAAAAGCCAGCCGGAGGCATCCCTGTGCTGGGCAGCCTGGTGAACACCGTCCTGAAGCACATC
ATCTGGCTGAAGGTCATCACAGCTAACATCCTCCAGCTGCAGGTGAAGCCCTCGGCCAATGACCA
GGAGCTGCTAGTCAAGATCCCCTGGACATGGTGGCTGGATTCAACACGCCCTGGTCAAGACCA
TCGTGGAGTTCACATGACGACTGAGGCCAAGCCACCATCCGCATGGACACCAGTGCAAGTGGC
CCCACCCGCTGGTCCTCAGTGACTGTGCCACCAGCCATGGGAGCCTGCGCATCCAACCTGCTGTA
TAAGCTCTCCTTCCTGGTGAACGCCTTAGCTAAGCAGGTGCATGAACCTCCTAGTGCCATCCCTGC
CCAATCTAGTGAAAAACCAGCTGTGTCCCGTGATCGAGGCTTCCTTCAATGGCATGTATGCAGAC
CTCCTGCAGCTGGTGAAGGTGCCATTTCCCTCAGCATTGACCGTCTGGAGTTGACCTTCTGTA
TCCTGCCATCAAGGGTGACACCATTACGCTCTACCTGGGGCCAAGTTGTTGGACTCACAGGGAA
AGGTGACCAAGTGGTTCAATAACTCTGCAGCTTCCCTGACAATGCCACCCTGGACAACATCCCG
TTCAGCCTCATCGTGAGTCAGGACGTGGTGAAGCTGCAGTGGCTGCTGTGCTCTCTCCAGAAGA
ATTCATGGTCCTGTTGGACTCTGTGCTTCCTGAGAGTGCCCATCGGCTGAAGTCAAGCATCGGGC
TGATCAATGAAAAGGCTGCAGATAAGCTGGGATCTACCCAGATCGTGAAGATCCTAACTCAGGAC
ACTCCCAGTTTTTTATAGACCAAGGCCATGCCAAGGTGGCCCAACTGATCGTGCTGGAAGTGT
TCCCTCCAGTGAAGCCCTCCGCCCTTTGTTACCCTGGGCATCGAAGCCAGCTCGGAAGCTCAGT
TTTACACCAAAGGTGACCAACTTATACTCAACTGAATAACATCAGCTCTGATCGGATCCAGCTG
ATGAACTCTGGGATTGGCTGGTTCCAACCTGATGTTCTGAAAACATCATCACTGAGATCATCCA
CTCCATCCTGCTGCCGAACCAGAATGGCAAATTAAGATCTGGGGTCCAGTGTCTATTGGTGAAGG
CCTTGGGATTCCGAGGCAGCTGAGTCTCACTGACCAAGGATGCCCTTGTGCTTACTCCAGCTCC
TTGTGGAACCCAGCTCTCCTGTCTCCAGTGAAGACTTGATGGCAGCCATCAGGGAAGGCTGG
GTCCAGCTGGGAGTATGGGTGTGAGCTCTATAGACCATCCCTCTCTGCAATCAATAAACACTG
CCTGTGAAAAA

FIGURE 78

MAGPWTFTLLCGLLAATLIQATLSPTAVLILGPKVIKEKLTQELKDNATSILQQPLLSAMREK
PAGGIPVLGSLVNTVLKHI IWLKVITANILQLQVKPSANDQELLVKIPLDMVAGFNTPLVKTI VE
FHMTTEAQATIRMDTSASGPTRLVLSDCATSHGSLRIQLLYKLSFLVNALAKQVMNLLVPSLPNL
VKNQLCPVIEASFNGMYADLLQLVKVPISLSIDRLEFDLLYPAIKGDTIQLYLGAKLLDSQGKVT
KWFNNSAASLTMP TLDNIPFSLIVSQDVVKA AVAVLSPEEFMVLLDSVLPESAHLKSSIGLIN
EKAADKLGSTQIVKILTQDTP EFFIDQGHAKVAQLIVLEVFPSEALRPLFTLGEASSEAQFYT
KGDQLILNLNNISSDRIQLMNSGIGWFQPDVLKNIITEI IHSILLPNQNGKLRSGVPVSLVKALG
FEAAESSLTKDALVLT PASLWKPSSPVSQ

Important features of the protein:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 48-51, 264-267, 401-404

Glycosaminoglycan attachment site.

amino acids 412-415

LBP / BPI / CETP family proteins.

amino acids 407-457

FIGURE 79

GAGAGAAGTCAGCCTGGCAGAGAGACTCTGAAATGAGGGATTAGAGGTGTTCAAGGAGCAAGAGC
TTCAGCCTGAAGACAAGGGAGCAGTCCCTGAAGACGCTTCTACTGAGAGGTCTGCCATGGCCCTCT
CTTGGCCTCCAACCTTGTGGGCTACATCCTAGGCCTTCTGGGGCTTTTGGGCACACTGGTTGCCAT
GCTGCTCCCCAGCTGGAAAACAAGTTCCTATGTGCGTGCCAGCATTGTGACAGCAGTTGGCTTCT
CCAAGGGCCTCTGGATGGAATGTGCCACACACAGCACAGGCATCACCCAGTGTGACATCTATAGC
ACCCCTTCTGGGCTGCCCGCTGACATCCAGGCTGCCCAGGCCATGATGGTGACATCCAGTGAAT
CTCCTCCCTGGCCTGCATTATCTCTGTGGTGGGCATGAGATGCACAGTCTTCTGCCAGGAATCCC
GAGCCAAAGACAGAGTGGCGGTAGCAGGTGGAGTCTTTTCATCCTTGAGGGCCTCCTGGGATTC
ATTCTGTGCTGGAATCTTCATGGGATCCTACGGGACTTCTACTCACCCTGGTGCCTGACAG
CATGAAATTTGAGATTGGAGAGGCTCTTACTTGGGCATTATTTCTTCCCTGTCTCCCTGATAG
CTGGAATCATCTCTGCTTTTCTGCTCATCCCAGAGAAATCGCTCCAACTACTACGATGCCTAC
CAAGCCCAACCTCTTGCCACAAGGAGCTCTCCAAGCCTGGTCAACCTCCCAAAGTCAAGAGTGA
GTTCAATCCTACAGCCTGACAGGGTATGTGTGAAGAACCAGGGGCCAGAGCTGGGGGGTGGCTG
GGTCTGTGAAAAACAGTGGACAGCACCCCGAGGGCCACAGGTGAGGGACACTACCCTGGATCGT
GTCAGAAGGTGCTGCTGAGGATAGACTGACTTTGGCCATTGGATTGAGCAAAGGCAGAAATGGGG
GCTAGTGTAACAGCATGCAGGTTGAATTGCCAAGGATGCTCGCCATGCCAGCCTTTCTGTTTTCC
TCACCTTGCTGCTCCCCTGCCCTAAGTCCCCAACCCCTCAACTTGAAACCCCATTCCTTAAGCCA
GGACTCAGAGGATCCCTTTGCCCTCTGGTTTACCTGGGACTCCATCCCCAACCCACTAATCACA
TCCCCTGACTGACCCTCTGTGATCAAAGACCCTCTCTCTGGCTGAGGTTGGCTCTTAGCTCATT
GCTGGGGATGGGAAGGAGAAGCAGTGGCTTTTGTGGGCATTGCTCTAACCTACTTCTCAAGCTTC
CCTCCAAAGAAACTGATTGGCCCTGGAACCTCCATCCCCTCTTGTTATGACTCCACAGTGTCCA
GACTAATTTGTGCATGAACTGAAATAAAACCATCCTACGGTATCCAGGGAACAGAAAGCAGGATG
CAGGATGGGAGGACAGGAAGGCAGCCTGGGACATTTAAAAAATA

FIGURE 8o

MASLGLQLVGYILGLLGLLGLTLVAMLLPSWKTSSYVGASIVTAVGFSKGLWMECATHSTGITQCD
IYSTLLGLPADIQAAQAMMVTSSAISSLACIISVVGMRCTVFCQESRAKDRVAVAGGVFFILGGL
LGFIPVAVNHLHGILRDFYSPLVPDSMKFEIGEALYLGIISSLFSLIAGIILCFSCSSQRNRSNY
DAYQAQPLATRSSPRPGQPPKVKSEFNYSYSLTGYV

Important features of the protein:

Signal peptide:

amino acids 1-24

Transmembrane domains:

amino acids 82-102, 117-140, 163-182

N-glycosylation site.

amino acids 190-193

PMP-22 / EMP / MP20 family proteins.

amino acids 46-59

FIGURE 81

CCCACGCGTCCGCGCCTCTCCCTTCTGCTGGACCTTCCTTCGTCTCTCCATCTCTCCCTCCTTTC
CCC GCGTTCCTTTCCACCTTTCTCTTCTTCCCACCTTAGACCTCCCTTCCTGCCCTCCTTTCCCT
GCCACCGCTGCTTCCTGGCCCTTCTCCGACCCCGTCTAGCAGCAGACCTCCTGGGGTCTGTGG
GTTGATCTGTGGCCCTGTGCCTCCGTGTCTTTTCGTCTCCCTTCCTCCCGACTCCGCTCCCGG
ACCAGCGGCCTGACCCTGGGAAAGGATGGTTCCCGAGGTGAGGGTCTCTCCTCCTTGCTGGGA
CTCGCGTGTCTGTGGTCCCCCTGGACTCCCACGCTCGAGCCCGCCAGACATGTTCTGCCTTTT
CCATGGGAAGAGATACTCCCCGGCGAGAGCTGGCACCCCTACTTGGAGCCACAAGGCCTGATGT
ACTGCCTGCGCTGTACCTGCTCAGAGGGCGCCCATGTGAGTTGTTACCGCCTCCACTGTCCGCT
GTCCACTGCCCCAGCCTGTGACGGAGCCACAGCAATGCTGTCCCAAGTGTGTGGAACCTCACAC
TCCCTCTGGACTCCGGGCCCCACAAAGTCTGCCAGCACAACGGGACCATGTACCAACACGGAG
AGATCTTCAGTGCCCATGAGCTGTTCCCTCCCGCTGCCAACCAGTGTGTCTCTGCAGCTGC
ACAGAGGGCCAGATCTACTGCGGCCTCACAACCTGCCCCGAACCAGGCTGCCAGCACCCCTCCC
ACTGCCAGACTCCTGCTGCCAAGCCTGCAAAGATGAGGCAAGTGAAGCAATCGGATGAAGAGGACA
GTGTGCAGTCGCTCCATGGGGTGTAGACATCCTCAGGATCCATGTTCCAGTGTGCTGGGAGAAAG
AGAGGCCCGGGCACCCAGCCCCACTGGCCTCAGCGCCCTCTGAGCTTCATCCCTCGCCACTT
CAGACCCAAGGGAGCAGGCAGCACAACCTGTCAAGATCGTCTGAAGGAGAAACATAAGAAAGCCT
GTGTGCATGGCGGGAAGACGTACTCCCACGGGGAGGTGTGGCACCCGGCCTTCCGTGCCTTCGGC
CCCTTGCCCTGCATCCTATGCACCTGTGAGGATGGCCGCCAGGACTGCCAGCGTGTGACCTGTCC
CACCGAGTACCCTGCCGTACCCCGAGAAAGTGGCTGGGAAGTGTGCAAGATTTGCCAGAGG
ACAAAGCAGACCCTGGCCACAGTGAGATCAGTTCTACCAGGTGTCCAAGGCACCGGGCCGGGTC
CTCGTCCACACATCGGTATCCCCAAGCCCAGACAACCTGCGTGCCTTTGCCCTGGAACACGAGGC
CTCGGACTTGGTGGAGATCTACCTCTGGAAGCTGGTAAAAGATGAGGAAACTGAGGCTCAGAGAG
GTGAAGTACCTGGCCCCAAGGCCACACAGCCAGAATCTTCCACTTGACTCAGATCAAGAAAGTCAG
GAAGCAAGACTTCCAGAAAGAGGCACAGCACTTCCGACTGCTCGCTGGCCCCACGAAGGTCAC
GGAACGTCTTCTAGCCCAGACCCTGGAGCTGAAGGTACGCGCCAGTCCAGACAAAGTGACCAAG
ACATAACAAAGACC TAACAGTTGCAGATATGAGCTGTATAATTGTTGTTATTATATATTAATAAA
TAAGAAGTTGCATTACCCTCAAAAAAAAAAAAAAAAAAAAAA

FIGURE 82

MVPEVRVLSLLGLALLWFPLDSHARARPDMFCLFHGKRYSPGESWHPYLEPQGLMYCLRCTCSE
GAHVSCYRLHCPPVHCPQPVTEPQCCPKVEPHTPSGLRAPPKSCQHNGTMYQHGEIFSAHELF
PSRLPNQCVLCSCTEGQIYCLTTCPEPGCPAPLPLPDSCCQACKDEASEQSDEEDSVQSLHGVR
HPQDPCSSDAGRKRGPPTPAPTGLSAPLSFIPRHFRPKGAGSTTVKIVLKEKHKKACVHGGKTYS
HGEVWHPAFRAFGPLPCILCTCEDGRQDCQRVTCPTTEYPCRHPKAVAGKCKICPEDKADPGHSE
ISSTRCPKAPGRVLVHTSVSPDNLRRFALEHEASDLVEIYLWKLVKDEETEAQRGEVPGPRPH
SQNLPLDSDQESQEARLPERGTALPTARWPPRRSLERLPSDPGAEGHGQSRQSDQDITKT

Signal peptide:

amino acids 1-25

FIGURE 83

GACAGCTGTGTCTCGATGGAGTAGACTCTCAGAACAGCGCAGTTTGCCCTCCGCTCACGCAGAGCCTCTCC
GTGGCTTCCGCACCTTGAGCATTAGGCCAGTTCTCCTTCTCTCTAATCCATCCGTACCTCTCCTGTCA
TCCGTTTCCATGCCGTGAGGTCCATTACAGAACACATCCATGGGCTCTCATGCTCAGTTTGGTTCTGAGTC
TCCTCAAGCTGGGATCAGGGCAGTGGCAGGTGTTGGGGCCAGACAAGCCTGTCCAGGCCTTGGTGGGGGAG
GACGCAGCATTCTCCTGTTTCCCTGTCTCCTAAGACCAATGCAGAGGCCATGGAAGTGCGGTTCTTCAGGGG
CCAGTTCCTAGCGTGGTCCACCTCTACAGGGACGGGAAGGACCAGCCATTTATGCAGATGCCACAGTATC
AAGGCAGGACAAAACCTGGTGAAGGATTCTATTGCGGAGGGGCGCATCTCTGAGGCTGGAAAACATTACT
GTGTTGGATGCTGGCCTCTATGGGTGCAGGATTAGTTCAGTCTTACTACCAGAAGGCCATCTGGGAGCT
ACAGGTGTCAGCACTGGGCTCAGTTCCTCTCATTTCATCACGGGATATGTTGATAGAGACATCCAGCTAC
TCTGTACGTCCTCGGGCTGGTTCCTCCCGGCCACAGCGAAGTGGAAGTCCACAAGGACAGGATTTGTCC
ACAGACTCCAGGACAAACAGAGACATGCATGGCCTGTTTGATGTGGAGATCTCTGACCCTCCAAGAGAA
CGCCGGGAGCATATCCTGTTCCATGCGGCATGCTCATCTGAGCCGAGAGGTGGAATCCAGGGTACAGATAG
GAGATACCTTTTTTCGAGCCTATATCGTGGCACCTGGCTACCAAAGTACTGGGAATACTCTGCTGTGGCCTA
TTTTTTGGCATTGTTGGACTGAAGATTTCTTCTCCAAATCCAGTGGAAAATCCAGGCGGAAGTGGACTG
GAGAAGAAAGCACGGACAGGCAGAATTGAGAGACGCCCGGAAACACGCAGTGGAGGTGACTCTGGATCCAG
AGACGGCTCACCGAAGCTCTGCGTTTCTGATCTGAAAAGTGTAAACCCATAGAAAAGCTCCCCAGGAGGTG
CCTCACTCTGAGAAGAGATTTACAAGGAAGAGTGTGGTGGCTTCTCAGAGTTTCCAAGCAGGGAACATTA
CTGGGAGGTGGACGGAGGACACAATAAAAGGTGGCGGTGGGAGTGTGCCGGATGATGTGGACAGGAGGA
AGGAGTACGTGACTTTGTCTCCCGATCATGGGTACTGGTCTCAGACTGAATGGAGAACATTTGTATTTCT
ACATTAATCCCGTTTTATCAGCGTCTTCCCAGGACCCACCTACAAAATAGGGGTCTTCTGGACTA
TGAGTGTGGGACCATCTCCTTCTTCAACATAAATGACCAGTCCCTTATTTATACCCTGACATGTCGGTTTG
AAGGCTTATTGAGGCCCTACATFGAGTATCCGTCTATAATGAGCAAAATGGAAGTCCCATAGTCATCTGC
CCAGTCAACCAGGAATCAGAGAAAGAGCCTCTTGGCAAAGGGCCTCTGCAATCCCAGAGACAAGCAACAG
TGAGTCTCTCACAGGCAACCAGCCCTTCTCCCCAGGGTGAAATGTAGGATGAATCACATCCCACAT
TCTTCTTTAGGGATATTAAGGTCTCTCTCCCAGATCCAAAGTCCCGCAGCAGCCGCCAAGGTGGCTTCCA
GATGAAGGGGGACTGGCCTGTCCACATGGGAGTCAGGTGTCATGGCTGCCCTGAGCTGGGAGGGAAGAAGG
CTGACATTACATTTAGTTGCTCTCACTCCATCTGGCTAAGTATCTTGAAATACCACCTCTCAGGTGAAG
AACCGTCAGGAATCCCATCTCACAGGCTGTGGTGTAGATTAAGTAGACAAGGAATGTGAATAATGCTTAG
ATCTTATTGATGACAGAGTGTATCCTAATGGTTTGTTCATTATATTACACTTTCAGTAAAAAA

FIGURE 84

MALMLSLVLSLLKLGSGQWQVFGPDKPVQALVGEDAAFCFLSPKTNAEAMEVRFFRGQFSSVVH
LYRDGKDQPFMQMPOYQGRTKLVKDSIAEGRISLRLENITVLDAGLYGCRISSQSYQKAIWELQ
VSALGSVPLISITGYVDRDIQLLCQSSGWFFRPRTAKWKGPQGDLSTDSRTNRDMHGLFDVEISL
TVQENAGSISCSMRHAHLSREVESRVQIGDTFFEPISWHLATKVLGILCCGLFFGIVGLKIFFSK
FQWKIQAELDWRRKHGQAEIRDARKHAVEVTLDPETAHPKLCVSDLKTVTHRKAQEVPHSEKRF
TRKSVVASQSFQAGKHYWEVDGGHNKRWRVGVCRDDVDRRKEYVTLSPDHGYYVLRRLNGEHLTYT
LNPRFISVFPRTPTTKIGVFLDYECGTISFFNINDQSLIYTLTCRFEGLLRPYIEYPSYNEQNGT
PIVICPVTQESEKEASWQRASAIPESTSSESSSQATTPFLPRGEM

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 239-255

FIGURE 85

AACAGACGTTCCCTCGCGGCCCTGGCACCTCTAACCCAGACATGCTGCTGCTGCTGCTGCCCT
GCTCTGGGGGAGGGAGAGGGCGGAAGGACAGACAAGTAAACTGCTGACGATGCAGAGTTCCGTGA
CGGTGCAGGAAGGCCTGTGTGCCATGTGCCCTGCTCCTTCTCCTACCCCTCGCATGGCTGGATT
TACCCTGGCCAGTAGTTTATGGCTACTGGTTCCGGGAAGGGGCCAATACAGACCAGGATGCTCC
AGTGGCCACAAACAACCCAGCTCGGGCAGTGTGGGAGGAGACTCGGGACCGATTCCACCTCCTTG
GGGACCCACATAACCAAGAATTGCACCCCTGAGCATCAGAGATGCCAGAAGAAGTGATGCGGGGAGA
TACTTCTTTTCGTATGGAGAAAGGAAGTATAAAATGGAATTATAAACATCACCGGCTCTCTGTGAA
TGTGACAGCCTTGACCCACAGGCCAACATCCTCATCCAGGCACCCCTGGAGTCCGGCTGCCCCC
AGAATCTGACCTGCTCTGTGCCCTGGGCCTGTGAGCAGGGGACACCCCTATGATCTCCTGGATA
GGGACCTCCGTGTCCCCCTGGACCCCTCCACCACCCGCTCCTCGGTGCTCACCCCTCATCCACA
GCCCCAGGACCATGGCACCAGCCTCACCTGTGAGGTGACCTTCCCTGGGGCCAGCGTGACCACGA
ACAAGACCGTCCATCTCAACGTGTCTACCCGCCTCAGAACTTGACCATGACTGTCTTCCAAGGA
GACGGCACAGTATCCACAGTCTTGGGAAATGGCTCATCTCTGTCACTCCAGAGGGCCAGTCTCT
GCGCCTGGTCTGTGCAGTTGATGCAGTTGACAGCAATCCCCCTGCCAGGCTGAGCCTGAGCTGGA
GAGGCCTGACCCTGTGCCCTCACAGCCCTCAAACCCGGGGTGTGGAGCTGCCTTGGGTGCAC
CTGAGGGATGCAGCTGAATTCACCTGCAGAGCTCAGAACCCTCTCGGCTCTCAGCAGGTCTACCT
GAACGTCTCCCTGCAGAGCAAAGCCACATCAGGAGTGACTCAGGGGTGGTCCGGGGAGCTGGAG
CCACAGCCCTGGTCTTCCCTGTCTTCTGCGTCATCTTCGTTGTAGTGAGGTCTGCAGGAAGAAA
TCGGCAAGGCCAGCAGCGGGCGTGGGAGATACGGGCATAGAGGATGCAAACGCTGTCAGGGGTTT
AGCCTCTCAGGGGCCCTGACTGAACCTTGGGCAGAAGACAGTCCCCAGACCAGCCTCCCCAG
CTTCTGCCCCCTCCTCAGTGGGGGAAGGAGAGCTCCAGTATGCATCCCTCAGCTTCCAGATGGTG
AAGCCTTGGGACTCGCGGGACAGGAGCCACTGACACCGAGTACTCGGAGATCAAGATCCACAG
ATGAAGAACTGCAGAGACTCACCCCTGATTGAGGGATCACAGCCCTCCAGGCAAGGGAGAAGTCA
GAGGCTGATTCTTGTAGAATTAACAGCCCTCAACGTGATGAGCTATGATAACACTATGAATTATG
TGCAGAGTGAAAAGCACACAGGCTTTAGAGTCAAAGTATCTCAAACCTGAATCCACACTGTGCC
TCCCTTTTATTTTTTAACTAAAAGACAGACAAATTCCTA

FIGURE 86

MLLLLLPLLWGRERAEGQTSKLLTMQSSVTVQEGLCVHVPCSFSYPSHGWIIYGPVHVHGYWFREG
ANTDQDAPVATNNPARAVWEE TRDRFHLLGDPHTKNCTLSIRDARRSDAGRYFFRMEKGSIKWNY
KHHRLSVNVTALTHRPNILIPGTLESGCPQNLTCVWPWACEQGTPPMISWIGTSVSPFLDPSTTRS
SVLTLPQPQDHGTSLTCQVTFPGASVTNKTVHLNVSYPQNLMTVFQGDGTVSTVLGNGSSL
SLPEGQSLRLVCAVDAVDSNPPARLSLSWRGLTLCPSQPSNPGVLELPWVHLRDAAEFTCRAQNP
LGSQQVYLNVS LQSKATSGVTQGVVGGAGATALVFLSFCVIFVVVRSRKRKSARPAAGVGDGTGIE
DANAVRGSASQGPLTEPWAEDSPPDQPPASARSSVGEQELQYASLSFQMVKPWDSRGQEA TDTE
YSEIKIHR

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 351-370

FIGURE 87

AGAAAGCTGCACTCTGTTGAGCTCCAGGGCCGAGTGGAGGGAGGGAGTGAAGGAGCTCTCTGTAC
CCAAGGAAAGTGCAGCTGAGACTCAGACAAGATTACAATGAACCAACTCAGCTTCCTGCTGTTTC
TCATAGCGACCACCAGAGGATGGAGTACAGATGAGGCTAATACTTACTTCAAGGAATGGACCTGT
TCTTCGTCTCCATCTCTGCCAGAAGCTGCAAGGAAATCAAAGACGAATGTCTTAGTGCAATTTGA
TGGCCTGTATTTCTCCGCACTGAGAATGGTGTATCTACCAGACCTTCTGTGACATGACCTCTG
GGGGTGGCGGTGGACCTGGTGGCCAGCGTGCATGAGAATGACATGCGTGGGAAGTGCACGGTG
GGCGATCGCTGGTCCAGTCAGCAGGGCAGCAAAGCAGACTACCCAGAGGGGGACGGCAACTGGGC
CAACTACAACACCTTTGGATCTGCAGAGGCGGCCACGAGCGATGACTACAAGAACCCTGGCTACT
ACGACATCCAGGCCAAGGACCTGGGCATCTGGCACGTGCCAATAAGTCCCCATGCAGCACTGG
AGAAACAGCTCCCTGCTGAGGTACCGCACGGACACTGGCTTCCTCCAGACACTGGGACATAATCT
GTTTGGCATCTACCAGAAATATCCAGTAAAATATGGAGAAGGAAAGTGTGGACTGACAACGGCC
CGGTGATCCCTGTGGTCTATGATTTTGGCGACGCCAGAAAACAGCATCTTATTACTCACCTAT
GGCCAGCGGAATTCACTGCGGGATTTGTTTCAGTTCAGGGTATTTAATAACGAGAGAGCAGCCAA
CGCCTTGTGTGCTGGAATGAGGGTCACCGGATGTAACACTGAGCATCACTGCATTGGTGGAGGAG
GATACTTCCAGAGGCCAGTCCCCAGCAGTGTGGAGATTTTCTGGTTTTGATTGGAGTGGATAT
GGAACTCATGTTGGTTACAGCAGCAGCCGTGAGATAACTGAGGCAGCTGTGCTTCTATTCTATCG
TTGAGAGTTTTGTGGGAGGGAACCCAGACCTCTCCPCCAACCATGAGATCCCAAGGATGGAGAA
CAACTTACCCAGTAGCTAGAATGTTAATGGCAGAAGAGAAAACAATAATCATATTGACTCAAGA
AAAAAA

FIGURE 88

MNQLSFLFLIATRSTDEANTYFKEWTCSSPSLPRSCKEIKDECPSAFDGLYFLRTENGVI
YQTFCDMTSGGGGWTILVASVHENDMRGKCTVGDRWSSQQGSKADYPEGDGNWANYNTFGSAEAAAT
SDDYKNPGYYDIQAKDLGIWHVPNKSPMQHWRNSSLLRYRTDTGFLQTLGHNLFGIYQKYPVKYG
EGKCWTDNGPVIPIVVYDFGDAQKTASYSPYQREFTAGFVQFRVENNERAANALCAGMRVTGCN
TEHHCIGGGGYFPEASPQQCGDFSGFDWSGYGTHVGYSSSREITEAAVLLFYR

Important features:

Signal peptide:

amino acids 1-16

N-glycosylation site.

amino acids 163-167

Glycosaminoglycan attachment sites.

amino acids 74-78, 289-293

N-myristoylation sites.

amino acids 76-82, 115-121, 124-130, 253-259, 292-298

FIGURE 89

CTAGATTTGTCGGCTTGC GGGGAGACTTCAGGAGTCGCTGTCTCTGAACTTCCAGCCTCAGAGAC
CGCCGCCCTTGTCCCCGAGGGCCATGGGCCGGGTCTCAGGGCTTGTGCCCTCTCGCTTCCCTGACG
CTCCTGGCGCATCTGGTGGTCGTCATCACCTTATTCTGGTCCCGGACAGCAACATACAGGCCTG
CCTGCCTCTCACGTTACCCCCGAGGAGTATGACAAGCAGGACATTCAGCTGGTGGCCGCGCTCT
CTGTCACCCTGGGCCTCTTTGCAGTGGAGCTGGCCGTTTCTCTCAGGAGTCTCCATGTTCAAC
AGCACCCAGAGCCTCATCTCCATTGGGGCTCACTGTAGTGCATCCGTGGCCCTGTCTTCTTCAT
ATTCGAGCGTTGGGAGTGCACCTACGTATTGGTACATTTTGTCTTCTGCAGTGCCCTTCCAGCTG
TCACTGAAATGGCTTTATTTCGTCACCGTCTTTGGGCTGAAAAGAAACCCTTCTGATTACCTCA
TGACGGGAACCTAAGGACGAAGCCTACAGGGGCAAGGGCCGCTTCGTATTCCTGGAAGAAGGAAG
GCATAGGCTTCGGTTTTCCCTCGGAAACTGCTTCTGCTGGAGGATATGTGTTGGAATAATTACG
TCTTGAGTCTGGGATTATCCGATTGTATTTAGTGCTTTGTAATAAAATATGTTTTGTAGTAAACA
TTAAGACTTATATACAGTTT TAGGGGACAATTA AAAAAAAAAA

FIGURE 91

CTGGGACCCCGAAAAGAGAAGGGGAGAGCGAGGGGACGAGAGCGGAGGAGGAAGATGCAACTGAC
TCGCTGCTGCTTCGTGTTCCCTGGTGCAGGGTAGCCTCTATCTGGTCATCTGTGGCCAGGATGATG
GTCTCCCGGGCTCAGAGGACCTGAGCGTGATGACCACGAGGGCCAGCCCCGGCCCCGGGTGCCT
CGGAAGCGGGGCCACATCTCACCTAAGTCCCGCCCCATGGCCAATCCACTCTCTAGGGCTGCT
GGCCCCGCTGGGGAGGCTTGGGGCATTCTTGGGCAGCCCCCAACCGCCGAACCACAGCCCCC
CACCTCAGCCAAGGTGAAGAAAATCTTTGGCTGGGGCGACTTCTACTCCAACATCAAGACGGTG
GCCTGAACCTGCTCGTCACAGGGAAGATTGTGGACCATGGCAATGGGACCTTCAGCGTCCACTT
CCAACACAATGCCACAGGCCAGGGAACATCTCCATCAGCCTCGTGCCCCCAGTAAAGCTGTAG
AGTTCACCAGGAACAGCAGATCTTCATCGAAGCCAAGGCCTCCAAAATCTTCAACTGCCGGATG
GAGTGGGAGAAGGTAGAACGGGGCCCGCGGACCTCGCTTTCACCCACGACCCAGCCAAGATCTG
CTCCCGAGACCAGCTCAGAGCTCAGCCACCTGGAGCTGCTCCCAGCCCTTCAAAGTCGTCTGTG
TCTACATCGCCTTCTACAGCACGGACTATCGGCTGGTCCAGAAGGTGTGCCAGATTACAACCTAC
CATAGTGATACCCCTACTACCCATCTGGGTGACCCCGGGCAGGCCACAGAGGCCAGGCCAGGGC
TGAAGGACAGGCCTGCCATGCAGGAGACCATCTGGACACCGGGCAGGGAAGGGGTGGGCCTC
AGGCAGGGAGGGGGTGGAGACGAGGAGATGCCAAGTGGGGCCAGGGCCAAGTCTCAAGTGGCAG
AGAAAGGGTCCCAAGTGCTGGTCCCAACCTGAAGCTGTGGAGTACTAGATCACAGGAGCACTGG
AGGAGGAGTGGGCTCTCTGTGCAGCCTCACAGGGCTTTGCCACGGAGCCACAGAGAGATGCTGGG
TCCCCGAGGCTGTGGGCAGGCCGATCAGTGTGGCCCCAGATCAAGTCATGGGAGGAAGCTAAGC
CCTTGGTTCTTGCCATCCTGAGGAAAGATAGCAACAGGGAGGGGAGATTTTCATCAGTGTGGACA
GCCTGTCAACTTAGGATGGATGGCTGAGAGGGCTTCCTAGGAGCCAGTCAGCAGGGTGGGGTGGG
GCCAGAGGAGCTCTCCAGCCCTGCCTAGTGGGGCCCTGAGCCCTTGTCTGTGCTGAGCATGG
CATGAGGCTGAAGTGGCAACCCTGGGGTCTTTGATGTCTTGACAGATTGACCATCTGTCTCCAGC
CAGGCCACCCCTTTCCAAAATTCCTCTTCTGCCAGTACTCCCTGTACCACCCATTGCTGATG
GCACACCCATCCTTAAGCTAAGACAGGACGATTGTGGTCTCCACACTAAGGCCACAGCCCATC
CGCGTGTGTGTCCCTCTTCCACCCCAACCCCTGCTGGCTCCTCTGGGAGCATCCATGTCCCG
GAGAGGGTCCCTCAACAGTCAGCCTCACCTGTGAGACGGGGTCTCCCGGATCTGGATGGCGC
CGCCCTCTCAGCAGCGGGCACGGGTGGGGCGGGCCGGCCGAGAGCATGTGCTGGATCTGTTT
TGTGTGTCTGTCTGTGGGTGGGGGAGGGGAGGGAAGTCTTGTGAAACCGCTGATTGCTGACTTT
TGTGTGAAGAATCGTGTCTTGGAGCAGGAAATAAAGCTTGCCCCGGGGCA

FIGURE 92

MQLTRCCFVFLVQGSlyLVICGQDDGPPGSEDPERDDHEGQPRPRVPRKRGHISPKSRPMANSTL
LGLLAPPGEAWGILGQPPNRPNHSPPPSAKVKKIFGWGDFYSNIKTVALNLLVTGKIVDHGNGTF
SVHFQHNATGQGNISISLVPPSKAVEFHQEQQIFIEAKASKIFNCRMWEKVERGRRTSLCTHDP
AKICSRDHAQSSATWCSQPFKVVCVYIAFYSTDYRLVQKVCPTYNYHSDTPYYPSG

Important features of the protein:

Signal peptide:

amino acids 1-14

N-glycosylation sites.

amino acids 62-65, 127-130, 137-140, 143-146

2-oxo acid dehydrogenases acyltransferase

amino acids 61-71

FIGURE 93

CGGTGGCCATGACTGCGGCCGTGTTCTTCGGCTGCGCCTTCATTGCCTTCGGGCCTGCGCTCGCC
CTTTATGTCTTCACCATCGCCATCGAGCCGTTGCGTATCATCTTCCTCATCGCCGGAGCTTCTT
CTGGTTGGTGTCTCTACTGATTTTCGTCCCTTGTTTGGTTCATGGCAAGAGTCATTATTGACAACA
AAGATGGACCAACACAGAAATATCTGCTGATCTTTGGAGCGTTTGTCTCTGTCTATATCCAAGAA
ATGTTCCGATTTGCATATTATAAACTCTTAAAAAAGCCAGTGAAGGTTTGAAGAGTATAAACCC
AGGTGAGACAGCACCTCTATGCGACTGCTGGCCTATGTTTCTGGCTTGGGCTTGGAAATCATGA
GTGGAGTATTTTCCTTTGTGAATACCTATCTGACTCCTTGGGGCCAGGCACAGTGGGCATTTCAT
GGAGATTCTCCTCAATTCTTCCTTTATTCAGCTTTCATGACGCTGGTCATTATCTTGCTGCATGT
ATTCTGGGGCATTGTATTTTTTGTATGGCTGTGAGAAGAAAAGTGGGGCATCCTCCTTATCGTTC
TCCTGACCCACCTGCTGGTGTGAGCCAGACCTTCATAAGTTCCTTATTATGGAATAAACCTGGCG
TCAGCATTATAATCCTGGTGTGCTGAGCCAGACCTTCATAAGTTCCTTATTATGGAATAAACCTGGCG
AAGCCTGAAACTCTGCCTGCTCTGCCAAGACAAGAACTTTCTTCTTTACAACCAGCGCTCCAGAT
AACCTCAGGGAACCAGCACTTCCCAAACCGCAGACTACATCTTTAGAGGAAGCACAACCTGTGCCT
TTTTCTGAAAATCCCTTTTTCTGGTGAATTGAGAAAGAAATAAACTATGCAGATA

FIGURE 94

MTAAVFFGCAFIAFGPALALYVFTIAIEPLRIIFLIAGAFFWLVSLLISSLVWFMARVIIDNKDG
PTQYLLIFGAFVSVYIQEMFRFAYYKLLKKASEGLKSNPGETAPSMRLLAYVSLGFGIMSGV
FSFVNTLSDSLGPVTGVIHGDSPQFFLYSAFMTLVIILLHVFWGIVFFDGCCKKKGILLIVLLT
HLLVSAQTFISSYYGINLASAFIILVLMGTWAFIAGGSCRSLKCLLCQDKNFLLYNQRSR

Important features of the protein:

Signal peptide:

amino acids 1-19

Transmembrane domains:

amino acids 32-51, 119-138, 152-169, 216-235

Glycosaminoglycan attachment site:

amino acids 120-123

Sodium:neurotransmitter symporter family protein

amino acids 31-65

FIGURE 95

AATTTTTCACCAGAGTAAACTTGAGAAACCAACTGGACCTTGAGTATTGTACATTTTGCCTCGTG
GACCCAAAGGTAGCAATCTGAAACATGAGGAGTACGATTCTACTGTTTTGTCTTCTAGGATCAAC
TCGGTCATTACCACAGCTCAAACCTGCTTTGGGACTCCCTCCCACAAAACCTGGCTCCGGATCAGG
GAACACTACCAAACCAACAGCAGTCAAATCAGGTCTTTCCCTTCTTTAAGTCTGATACCATTAACA
CAGATGCTCACACTGGGGCCAGATCTGCATCTGTAAATCCTGCTGCAGGAATGACACCTGGTAC
CCAGACCCACCCATTGACCTGGGAGGGTTGAATGTACAACAGCAACTGCACCCACATGTGTTAC
CAATTTTTGTACACAACCTGGAGCCCAGGGCACTATCCTAAGCTCAGAGGAATTGCCACAAATC
TTCACGAGCCTCATCATCCATTCTTGTTCCTGGGAGGCATCCTGCCACCAGTCAGGCAGGGGC
TAATCCAGATGTCCAGGATGGAAGCCTTCCAGCAGGAGGAGCAGGTGTAATCCTGCCACCCAGG
GAACCCAGCAGGCCGCCTCCCAACTCCCAGTGGCACAGATGACGACTTTGCAGTGACCACCCCT
GCAGGCATCCAAAGGAGCACACATGCCATCGAGGAAGCCACCACAGAATCAGCAAATGGAATTCA
GTAAGCTGTTTCAAATTTTTTCAACTAAGCTGCCTCGAATTTGGTGATACATGTGAATCTTTATC
ATTGATTATATTATGGAATAGATTGAGACACATTGGATAGTCTTAGAAGAAATTAATTCTTAATT
TACCTGAAAATATTCTTGAATTTTCAGAAAATATGTTCTATGTAGAGAATCCCAACTTTTAAAAA
CAATAATTC AATGGATAAATCTGTCTTTGAAATATAACATTATGCTGCCTGGATGATATGCATAT
TAAAACATATTTGAAAACCTGGAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 96

MRSTILLFCLLGSTRSLPQLKPALGLPPTKLAPDQGTLPNQOQSNQVFPSSLIPLTQM
LTLGPDHLHLLNPAAGMTPGTQTHPLTLGGLNVQQQLHPHVLPIFVTQLGAQGTILSSEE
LPQIFTSLLIIHSLFPGGILPTSQAGANPDVQDGSLPAGGAGVNPATQGTTPAGRLPTPSG
TDDDFAVTTPAGIQRSTHAIIEEATTESANGIQ

Signal peptide:

amino acids 1-16

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

FIGURE 98

MVPAWLWLLCVSVPQALPKAQAELSVEVPENYGGNFPLYLTKLPLPREGAEGQIVLSGDSGKAT
EGPFAMDPDSGFLLVTRALDREEQAEYQLQVTLEMQDGHVVLWGPQPVLVHVKDENDQVPHFSQAI
YRARLSRGTRPGIPFLFLEASDRDEPGTANSDLRFHILSQAPAQSPDMFQLEPRLGALALSPKG
STSLDHALERTYQLLVQVKMDQASGHQATATVEVSI IESTWVSLEPIHLAENLKVLYPHMAQ
VHWSGGDVHYHLESHPPGPFVNAEGNLYVTRELDREAQAEYLLQVRAQNSHGEDIAAPLELHVL
VMDENDNVPICPPRPDPTVSIPELSPPGTEVTRLSAEDADAPGSPNSHVVYQLLSPEPEDGVEGRA
FQVDPTSGSVTLGVLPLRAGQNILLV LAMDLAEGGFSSTCEVEVAVTDINDHAPEFITSQIG
PISLPEDVEPGTLVAMLTADADLEPAFRIMDFAIERGDTEGTFGLDWEPPDSGHVRLRLCKNLSY
EAAPSHEVVVVQSVAKLVGPGPGGATATVTVLVERVMPPPKLDQESYEASVPI SAPAGSFLLT
IQPSDPI SRTLRFSLVNDSEGWLCIEKFSGEVHTAQSLQGAQPGDTYTVLVEAQDTALT LAPVPS
QYLCTPRQDHGLIVSGPSKDPDLASGHGPYSFTLGNPTVQRDWRQLQTLNGSHAYLTALHWVEP
REHIIPVVVSHNAQMWQLLVRVIVCRNVEGQCMRKVGRMKGMPTKLSAVGIIVGTLVAIGIFLI
LIFTHWTMSRKKDPDQPADSVPLKATV

Signal peptide:

amino acids 1-18

Transmembrane domain:

amino acids 762-784

FIGURE 99

GGCTGACCGTGCTACATTGCTGGAGGAAGCCTAAGGAACCCAGGCATCCAGCTGCCACGCCTG
AGTCCAAGATTCTTCCCAGGAACACAAACGTAGGAGACCCACGCTCCTGGAAGCACCAGCCTTTA
TCTCTTACCTTCAAGTCCCCTTTCTCAAGAATCCTCTGTTCTTTGCCCTCTAAAGTCTTGGTAC
ATCTAGGACCCAGGCATCTTCTTTCCAGCCACAAAGAGACAGATGAAGATGCAGAAAGGAAATG
TTCTCCTTATGTTTGGTCTACTATTGCATTTAGAAGCTGCAACAAATTTCCAATGAGACTAGCACC
TCTGCCAACACTGGATCCAGTGTGATCTCCAGTGGAGCCAGCACAGCCACCAACTCTGGGTCCAG
TGTGACCTCCAGTGGGTCAGCACAGCCACCATCTCAGGGTCCAGCGTGACCTCCAATGGGGTCA
GCATAGTCACCAACTCTGAGTTCATACACCTCCAGTGGGATCAGCACAGCCACCAACTCTGAG
TTCAGCACAGCGTCCAGTGGGATCAGCATAGCCACCAACTCTGAGTCCAGCACAACTCCAGTGG
GCCAGCACAGCCACCAACTCTGAGTCCAGCACACCTCCAGTGGGGCCAGCACAGTCCACCAACT
CTGGGTCCAGTGTGACCTCCAGTGGAGCCAGCACTGCCACCAACTCTGAGTCCAGCACAGTGTCC
AGTAGGGCCAGCACTGCCACCAACTCTGAGTCTAGCACACTCTCCAGTGGGGCCAGCACAGCCAC
CAACTCTGACTCCAGCACAACTCCAGTGGGGCTAGCACAGCCACCAACTCTGAGTCCAGCACAA
CCTCCAGTGGGGCCAGCCACCAACTCTGAGTCCAGCACAGTGTCCAGTAGGGCCAGCACT
GCCACCAACTCTGAGTCCAGCACAACTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCCAG
AACGACCTCCAATGGGGTGGCACAGCCACCAACTCTGAGTCCAGCACAGCACTCCAGTGGGGCCA
GCACAGCCACCAACTCTGACTCCAGCACAGTGTCCAGTGGGGCCAGCACTGCCACCAACTCTGAG
TCCAGCACGACCTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACGACCTCCAGTGG
GGTAGCACAGCCACCAACTCTGACTCCAGCACAACTCCAGTGGGGCCGGCACAGCCACCAACT
CTGAGTCCAGCACAGTGTCCAGTGGGATCAGCACAGTCCACCAATCTGAGTCCAGCACACCTCC
AGTGGGGCCAACACAGCCACCAACTCTGAGTCCAGTACGACCTCCAGTGGGGCCAACACAGCCAC
CAACTCTGAGTCCAGCACAGTGTCCAGTGGGGCCAGCACTGCCACCAACTCTGAGTCCAGCACAA
CCTCCAGTGGGGTCCAGCACAGCCACCAACTCTGAGTCCAGCACAACTCCAGTGGGGCTAGCACA
GCCACCAACTCTGACTCCAGCACAACTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCTAG
CACAGTGTCCAGTGGGATCAGCACAGTCCACCAATCTGAGTCCAGCACAACTCCAGTGGGGCCA
ACACAGCCACCAACTCTGGGTCCAGTGTGACCTCTGCAGGCTCTGGAACAGCAGCTCTGACTGGA
ATGCACACAACTTCCCATAGTGCATCTACTGCAGTGTGAGGCAAAGCCTGGTGGGTCCCTGGT
GCCGTGGGAAATCTTCTCATCACCTGGTCTCGGTGTGGCGCCGTGGGGCTCTTTGCTGGGC
TCTTCTTCTGTGTGAGAAACAGCCTGTCCCTGAGAAACACCTTTAACACAGCTGTCTACCACCT
CATGGCTCAACCATGGCCTTGGTCCAGGCCCTGGAGGGAATCATGGAGCCCCCACAGGCCAG
GTGGAGTCTAATGTTCTGGAGGAGACCAGTATCATCGATAGCCATGGAGATGAGCGGGAGGA
ACAGCGGGCCCTGAGCAGCCCCGGAAGCAAGTCCCGCATTTCTCAGGAAGGAAGAGACCTGGGCA
CCCAAGACCTGGTTTCCTTTCATTCATCCAGGAGACCCCTCCAGCTTTGTTTGGATCCTGAA
AATCTTGAAGAAGGTATCTCACCTTTCTTGCCTTTACCAGACACTGGAAAGAGAATACTATAT
TGCTCATTTAGCTAAGAAATAAATACATCTCATCTAACACACACGACAAAGAGAAGCTGTGCTTG
CCCCGGGTGGGTATCTAGCTCTGAGATGAACTCAGTTATAGGAGAAAACCTCCATGCTGGACTC
CATCTGGCATTCAAAATCTCCACAGTAAAATCCAAAGACCTCAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

10063728

FIGURE 101

GGCCGGACGCCTCCGCGTTACGGGATGAATTAACGGCGGGTTCCGCACGGAGGTTGTGACCCCTA
CGGAGCCCCAGCTTGCCACGCACCCCACTCGGCGTCGCGCGGCGTGCCCTGCTTGTACAGGTG
GGAGGCTGGAACATCAGGCTGAAAAACAGAGTGGGTACTCTCTTCTGGAAGCTGGCAACAAAT
GGATGATGTGATATATGCATTCCAGGGGAAGGGAAATTGTGGTGCTTCTGAACCCATGGTCAATT
AACGAGGCAGTTTCTAGCTACTGCACGTACTTCATAAAGCAGGACTCTAAAAGCTTTGGAATCAT
GGTGTATGGAAAGGGATTTACTTTATACTGACTCTGTTTTGGGGAAGCTTTTTTGGAAAGCATTT
TCATGCTGAGTCCCTTTTTACCTTTGATGTTTGTAACCCATCTTGGTATCGCTGGATCAACAAC
CGCCTTGTGGCAACATGGCTCACCTACCTGTGGCATTATTGGAGACCATGTTTGGTGTAAAAGT
GATTATAACTGGGGATGCATTTGTTCCCTGGAGAAAGAAGTGTCAATTATCATGAACCATCGGACAA
GAATGGACTGGATGTTCCCTGTGGAATTGCTGATGCGATATAGCTACCTCAGATTGGAGAAAATT
TGCCTCAAAGCGAGTCTCAAAGGTGTTCCCTGGATTTGGTTGGGCCATGCAGGCTGCTGCCTATAT
CTTCATTCATAGGAAATGGAAGGATGACAAGAGCCATTTCAAGACATGATTGATTACTTTTGTG
ATATTCACGAACCACTTCAACTCCTCATATTTCCAGAAAGGGACTGATCTCACAGAAAACAGCAAG
TCTCGAAGTAATGCATTTGCTGAAAAAATGGACTTCAGAAATATGAATATGTTTTACATCCAAG
AACTACAGGCTTTACTTTTTGTGGTAGACCGTCTAAGAGAAGGTAAGAACCTTGATGCTGTCCATG
ATATCACTGTGGCGTATCCTCACAACATTCCTCAATCAGAGAAGCACCTCCTCCAAGGAGACTTT
CCCAGGGAAATCCACTTTTACGTCACCGGTATCCAATAGACACCTCCCACATCCAAGGAGGA
CCTTCAACTCTGGTGCCACAAACGGTGGGAAGAGAAAGAAGAGAGGCTGCGTTCCTTCTATCAAG
GGGAGAAGAATTTTTATTTTACCGGACAGAGTGTCAATTCACCTTGCAAGTCTGAACTCAGGGTC
CTTGTGGTCAAATTGCTCTCTATACTGTATTGGACCCTGTTTACGCCCTGCAATGTGCCTACTCAT
ATATTTGTACAGTCTTGTAAAGTGGTATTTTATAATCACCATTGTAATCTTTGTGCTGCAAGAGA
GAATATTTGGTGGACTGGAGATCATAGAACTTGCATGTTACCGACTTTTACACAAAACAGCCACAT
TTAAATTCAAAAGAAAATGAGTAAGATTATAAGTTTGCCATGTGAAAACCTAGAGCATATTTG
GAAATGTTCTAAACCTTTCTAAGCTCAGATGCATTTTGCATGACTATGTGGAATATTTCTTACT
GCCATCATATTTGTTAAAGATATTTGCACTTAATTTTGTGGAAAAATATTGCTACAATTTTT
TTTAATCTCTGAATGTAATTTGATACTGTGTACATAGCAGGGAGTGATCGGGGTGAAATAACTT
GGCCAGAATATTATTAACAATCATCAGGCTTTTAAA

FIGURE 102

MHSRGREIVVLLNPWSINEAVSSYCTYFIKQDSKSGIMVSWKGIYFILTLFWGSFFGSI FMLS P
FLPLMFVNPSWYRWINNRLVATWLTLPVALLETMFGVKVIITGDADFVPGERSVIIMNHRTRMDWM
FLWNC LMRYSYLRLEKICLKASLKGVPGF GWAMQAAAYIFIHRKWKDDKSHFEDMIDYFCDIHEP
LQLLIFPEGTDLTENSRSR SNAFAEKNGLQKYEYVLHPRTTGFTFVVDRLREGKNLDAVHDI TVA
YPHNIPQSEKHL LQGDFPREIHFHVHRYPIDTLPTS KEDLQLWCHKRWEEKEERLRSFYQGEKNF
YFTGQSVIPPCKSELRVLVVKLLSILYWTLFSPAMCLLIYLYSLVKWYFIITIVIFVLQERIFGG
LEIIE LACYRLLHKQPHLNSKKNE

Important features of the protein:

Signal peptide:

amino acids 1-22

Transmembrane domains:

amino acids 44-63, 90-108, 354-377

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400

FIGURE 104

MAFVLILVLSFYELVSGQWQVTGPGKFVQALVGEDAVFSCSLFPETSAEAMEVRFNRNQFHAVVH
LYRDGEDWESKQMPQYRGRTEFVKDSIAGGRVSLRLKNITPSDIGLYGCWFSSQIYDEEATWELR
VAALGSLPLISIVGYVDGGIQLLCLSSGWFPQPTAKWKGPQGDLSSDSRANADGYSLYDVEISI
IVQENAGSILCSIHLAEQSHEVESKVLIGETFFQPSWRLASILLGLLCGALCGVVMGMIIVFFK
SKGKIQAELDWRKKGQAEIRDARKHAVEVTLDPETAHPKLCVSDLKTVTHRKAQEVPHSEKRF
TRKSVVASQGFQAGRHYWEVDVGQNVGWYVGVCRDDVDRGKNNVTLSPNNGYWVLRLLTTEHLYFT
FNPHFISLPPSTPPTRVGVFLDYEGGTISFFNTNDQSLIYTLTLLTCQFEGLLRPYIQHAMYDEEKG
TPIFICPVSWG

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 131-150, 235-259

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300

FIGURE 105

CCTTCACAGGACTCTTCATTGCTGGTTGGCAATGATGTATCGGCCAGATGTGGTGAGGGCTAGGAAAAGAG
TTTGGTTGGGAACCTGGGTTATCGGCCTCGTCATCTTCATATCCCTGATTGTCTGGCAGTGTGCATTGGA
CTCACTGTTTATTATGTGAGATATAATCAAAGAAGACCTACAATTACTATAGCACATTGTCAATTTACAAC
TGACAACTATATGCTGAGTTTGGCAGAGAGGGCTTCTAACAATTTTACAGAAATGAGCCAGAGACTTGAAT
CAATGGTGAAAAATGCATTTTATAAATCTCCATTAAGGGAAGAATTTGTCAAGTCTCAGGTTATCAAGTTC
AGTCAACAGAAGCATGGAGTGTGGCTCATATGCTGTTGATTTGTAGATTTCACTCTACTGAGGATCCTGA
AACTGTAGATAAAAATTTTCAACTTGTTTTACATGAAAAGCTGCAAGATGCTGTAGGACCCCTAAAGTAG
ATCCTCACTCAGTTAAAATTAATAAATCAACAAGACAGAAACAGACAGCTATCTAAACCATTGCTGCGGA
ACACGAAGAAGTAAAATCTAGGTCAAGTCTCAGGATCGTTGGTGGGACAGAAGTAGAAGAGGGTGAATG
GCCCTGGCAGGCTAGCCTGCAGTGGGATGGGAGTCATCGCTGTGGAGCAACCTTAATTAATGCCACATGGC
TTGTGAGTGTCTCACTGTTTACAACATATAAGAACCCTGCCAGATGGACTGCTTCCTTTGGAGTAACA
ATAAAACCTTCGAAAATGAAACGGGGTCTCCGGAGAATAATTGTCCATGAAAAATACAAACACCCATCACA
TGACTATGATATTTCTCTTGCAGAGCTTCTAGCCCTGTTCCCTACACAAATGCAGTACATAGAGTTTGTCT
TCCCTGATGCATCCTATGAGTTTCAACCAGGTGATGTGATGTTTGTGACAGGATTTGGAGCACTGAAAAAT
GATGGTTACAGTCAAATCATCTTCGACAAGCACAGGTGACTCTCATAGACGCTACAACCTGCAATGAACC
TCAAGCTTACAATGACCCATAACTCCTAGAATGTTATGTGCTGGCTCCTTAGAAGGAAAAACAGATGCAT
GCCAGGGTACTCTGGAGGACCACTGGTTAGTTCAGATGCTAGAGATATCTGGTACCTTGCTGGAAATAGTG
AGCTGGGGAGATGAATGTGCGAAACCAACAAGCCTGGTGGTTTATACTAGAGTTACGGCCTTGCAGGACTG
GATTACTTCAAAAACCTGGTATCTAAGAGACAAAAGCCTCATGGAACAGATAACATTTTTTTTTGTTTTTGT
GGTGTGGAGGCCATTTTAGAGATACAGAATTGGAGAAGACTTGCAAAACAGCTAGATTTGACTGATCTCA
ATAAATGTTTGGCTTGATGCATGATTTTCTTCCAGCTCTGTTCCGCACGTAAAGCATCCTGCTTCTGCCA
GATCAACTCTGTCTGTGAGCAATAGTTGAAACTTTATGTACATAGAGAAATAGATAATACAATATTAC
ATTACAGCCTGTATTCAATTTGTTCTCTAGAAGTTTGTGAGAAATTTGACTTGTGACATAAATTTGTAAT
GCATATATACAATTTGAAGCACTCCTTTTCTTCAAGTTCCTCAGCTCCTCATTTCAGCAAATATCCATTT
TCAAGGTGCAGAACAGGAGTGAAAGAAAATATAAGAAGAAAAAATCCCCTACATTTTATTGGCACAGAA
AAGTATTAGGTGTTTTCTTAGTGGAAATATTAGAATGATCATATTCATTATGAAAGTCAAGCAAAGACA
GCAGAATACCAATCACTTCATCATTTAGGAAGTATGGGAAGTAAAGTAAAGGAAGTCCAGAAAGCAAGCAAG
ATATATCCTTATTTTCATTTCCAAACAACACTACTATGATAAATGTGAAGAAGATTCTGTTTTTTTGTGACCT
ATAATAATTATACAACTTCATGCAATGTACTTGTCTAAGCAAATTAAGCAAATATTTATTTAACATTG
TTACTGAGGATGTCAACATATAACAATAAAATATAAATCACCCA

FIGURE 106

MMYRPDVVRARKRVCWEPWVIGLVI F I S L I V L A V C I G L T V H Y V R Y N Q K K T Y N Y S T L S F T T D K L Y
A E F G R E A S N N F T E M S O R L E S M V K N A F Y K S P L R E E F V K S Q V I K F S Q Q K H G V L A H M L L I C R F H S T E D
P E T V D K I V Q L V L H E K L Q D A V G P P K V D P H S V K I K K I N K T E T D S Y L N H C C G T R R S K T L G Q S L R I V G G
T E V E E G E W P W Q A S L Q W D G S H R C G A T L I N A T W L V S A A H C F T T Y K N P A R W T A S F G V T I K P S K M K R G L
R R I V H E K Y K H P S H D Y D I S L A E L S S P V P Y T N A V H R V C L P D A S Y E F Q P G D V M F V T G F G A L K N D G Y S
Q N H L R Q A Q V T L I D A T T C N E P Q A Y N D A I T P R M L C A G S L E G K T D A C Q G D S G G P L V S S D A R D I W Y L A G
I V S W G D E C A K P N K P G V Y T R V T A L R D W I T S K T G I

Transmembrane domain:

amino acids 21-40 (type II)

FIGURE 107

AGAGAAAGAAGCGTCTCCAGCTGAAGCCAATGCAGCCCTCCGGCTCTCCGCGAAGAAGTTCCCTG
CCCCGATGAGCCCCCGCGTCCCGGACTATCCCCAGGCGGGCGTGGGGCACCAGGGCCCAGC
GCCGACGATCGTCCCGTTTTGCCCTTGGGAGTAGGATGTGGTGAAGGATGGGGCTTCTCCCTT
ACGGGGCTCACAATGCGCCAGAGAAGATTCCGTGAAGTGTCTGCGCTGCCTGCTCTACGCCCTCAA
TCTGCTCTTTTGGTTAATGTCCATCAGTGTGTTGGCAGTTTCTGCTTGGATGAGGGACTACCTAA
ATAATGTTCTCACTTTAACTGCAGAAACGAGGGTAGAGGAAGCAGTCATTTGACTTACTTTCTT
GTGGTTTCATCCGGTCATGATTGCTGTTTCTGCTGTTTCTTATCATTGTTGGGATGTTAGGATTTG
TGGAACGGTGAAGAATACTGTTGCTTCTGTCATGGTACTTTGGAAGTTTGGCTTGTCAATTTTCT
GTGTAGAACTGGCTTGTGGCGTTTGGACATATGAACAGGAACTTATGGTTCCAGTACAATGGTCA
GATATGGTCACTTTGAAGCCAGGATGACAAATTATGGATTACCTAGATATCGGTGGCTTACTCA
TGCTTGGAAATTTTTTTCAGAGAGAGTTTTAAGTGTCTGTGGAGTAGTATATTTCACTGACTGGTTGG
AAATGACAGAGATGGACTGGCCCCAGATTCCTGCTGTGTTAGAGAATTTCCAGGATGTTCCAAA
CAGGCCACCAGGAAGATCTCAGTGACCTTTATCAAGAGGGTTGTGGGAAGAAAATGTATTCTCTT
TTGAGAGGAACCAACAACACTGCAGGTGCTGAGGTTTTCTGGGAATCTCCATTGGGGTGACACAAA
TCCTGGCCATGATTCTCACCATTACTCTGCTCTGGGCTCTGTATTATGATAGAAGGGAGCCTGGG
ACAGACCAATGATGTCCTTGAAGAATGACAACCTCTCAGCACCTGTCATGTCCCTCAGTAGAACT
GTTGAAACCAAGCCTGTCAAGAATCTTTGAACACACATCCATGGCAAACAGCTTTAATACACACT
TTGAGATGGAGGAGTTATAAAAAGAAATGTCACAGAAGAAAACCACAAACTTGTTTTATTGGACT
TGTGAATTTTGTAGTACATACTATGTGTTTTCAGAAATATGTAGAAATAAAAATGTTGCCATAAAA
TAACACCTAAGCATATACTATTCTATGCTTTAAAATGAGGATGGAAAAGTTTTCATGTCATAAGTC
ACCACCTGGACAATAATTGATGCCCTTAAAATGCTGAAGACAGATGTCATACCCACTGTGTAGCC
TGTGTATGACTTTTACTGAACACAGTTATGTTTTGAGGCAGCATGGTTTGATTAGCATTTCGCCA
TCCATGCAAACGAGTCACATATGGTGGGACTGGAGCCATAGTAAAGGTTGATTTACTTCTACCAA
CTAGTATATAAAGTACTAATAAATGCTAACATAGGAAGTTAGAAAATACTAATAACTTTTATTA
CTCAGCGATCTATTCTCTGATGCTAAATAAATTATATATCAGAAAACCTTTCATATTGGTGACT
ACCTAAATGTGATTTTTGCTGGTTACTAAAATATTTCTACCCTTAAAAGAGCAAGCTAACACAT
TGTCTTAAGCTGATCAGGGATTTTTGTATATAAGTCTGTGTTAAATCTGTATAATTCAGTCGAT
TTCAGTTCTGATAATGTTAAGAATAACCATTATGAAAAGGAAAATTTGTCCTGTATAGCATCATT
ATTTTTAGCCTTTCCTGTTAATAAAGCTTTACTATTCTGTCTGGGCTTATATTACACATATAAC
TGTATTTTAAATACTTAACCACTAATTTTGAATAATACCAGTGTGATACATAGGAATCATTATTC
AGAATGTAGTCTGGTCTTTAGGAAGTATTAATAAGAAAATTTGCACATAACTTAGTTGATTGAGA
AAGGACTGTATGCTGTTTTTCTCCAAATGAAGACTCTTTTTGACACTAAACACTTTTTAAAAA
GCTTATCTTTGCCTTCTCCAACAAGAAGCAATAGTCTCCAAGTCAATATAAATCTACAGAAAA
TAGTGTCTTTTTCTCCAGAAAATGCTTGTGAGAATCATTAAAACATGTGACAATTTAGAGATT
CTTTGTTTTATTTCACTGATTAATATACTGTGGCAAATTACACAGATTATTAATTTTTTTACAA
GAGTATAGTATATTTTAAATGGGAAAAGTGCATTTTACTGTATTTTGTGATTTTGTGTTTAT
TTCTCAGAAATATGAAAGAAAATTAATAATGTGTCAATAAATATTTTCTAGAGAGTAA

FIGURE 108

MAREDSVKCLRCLLYALNLLFWLMSISVLAVSAWMDYLNNVLTTLAETRVEEAVILTYFPVVHP
VMIAVCCFLIIVGMLGYCGTVKRNLLLLAWYFGSLLVIFCVELACGVWVYEQELMVPVQWSDMVT
LKARMTNYGLPRYRWLTHAWNFFQREFKCCGVVYFTDWLEMTEMDWPPDSCCVREFPGCSKQAHQ
EDLSDLYQEGCGKMYSLRGTQQLQVLRFLGISIGVTQILAMILTITLLWALYDRREPGTDQM
MSLKNDNSQHLSCPSVELLKPSLSRI FEHTSMANSFNTHFEMEEL

Signal peptide:

amino acids 1-33

Transmembrane domains:

amino acids 12-35, 57-86, 94-114, 226-248

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

FIGURE 109

CCAAGGCCAGAGCTGTGGACACCTTATCCCACCTCATCCTCTTCCCTCTGATAAAGCCCCCTACCAGTGTCT
GATAAAGTCTTTCTCGTGAGAGCCTAGAGGCCCTAAAAAAGTGTCTTGAAGAGAAGGGGACAAAGGAACA
CCAGTATTAAGAGGATTTCCAGTGTCTTGGCAGTTGGTCCAGAAGGATGCCTCCATTCTGCTTCTCACCTG
CCTCTTCATCACAGGCACCTCCGTGTCAACCCGTGGCCCTAGATCCTTGTCTGCTTACATCAGCCTGAATGAGC
CCTGGAGGAACACTGACCACCAGTTGGATGAGTCTCAAGGTCCCTCCTATGTGACAACCATGTGAATGGGGAG
TGGTACCACCTTACGGGCATGGCGGGAGATGCCATGCCTACCTTCTGCATACCAGAAAACCACTGTGGAACCCA
CGCACCTGTCTGGCTCAATGGCAGCCACCCCTAGAAGGCGACGGCATTGTGCAACGCCAGGCTTGTGCCAGCT
TCAATGGAACTGCTGTCTCTGGAACACCACGGTGGAAAGTCAAGGCTTGGCCCTGGAGGCTACTATGTGTATCGT
CTGACCAAGCCCAGCGTCTGCTTCCACGTCTACTGTGGTCATTTTATGACATCTGCGACGAGGACTGCCATGG
CAGCTGCTCAGATAACCAGCGAGTGCACATGCCTCCAGGAACTGTGCTAGGCCCTGACAGGCAGACATGCTTTG
ATGAAAATGAATGTGAGCAAACACGGTGGCTGCAGTGAGATCTGTGTGAACCTCAAAAACCTCCTACCCTGT
GAGTGTGGGTTGGCCGTGTGCTAAGAAGTGTGGCAAGACTTGTGAAGACGTTGAAGGATGCCACAATAACAA
TGGTGGCTGCAGCCACTCTTGCCTTGGATCTGAGAAAGGCTACCAGTGTGAATGTCCCCGGGGCCTGGTGTGT
CTGAGGATAACCACACTTGCCTAAGTCCCTGTGTGTGCAAATCAAATGCCATTGAAGTGAACATCCCCAGGGAG
CTGGTTGGTGGCCTGGAGCTCTTCTGACCAACACCTCCTGCCGAGGAGTGTCCAACGGCACCCATGTCAACAT
CCTCTCTCTCTCAAGACATGTGGTACAGTGGTCGATGTGGTGAATGACAAGATTGTGGCCAGCACCTCGTGA
CAGGCTTACCAAGCAGACCCCCGGGAGCAGCGGGACTTCATCATCCGAACCAGCAAGCTGCTGATCCCGGTG
ACCTGCGAGTTTCCACGCTGTACACCATTTCTGAAGGATACGTTCCCAACCTTCGAAACTCCCCACTGGAAT
CATGAGCCGAAATCATGGGATCTTCCCATTCACTCTGGAGATCTTCAAGGACAATGAGTTTGAAGAGCCTTACC
GGGAAGCTCTGCCACCCCTCAAGCTTGGTACTCCCTCTACTTTGGCATTGAGCCCGTGGTGCACGTGAGCGGC
TTGGAAGCTTGGTGGAGAGCTGCTTTGCCACCCCACTCCAAGATCGACGAGGTCTGAAATACTACCTCAT
CCGGGATGGCTGTGTTTCAGATGACTCGGTAAGCAGTACACATCCCGGGATCACCTAGCAAAGCACTTCCAGG
TCCCTGTCTTCAAGTTTGTGGGCAAAGACCACAAGGAAGTGTCTGCACTGCCGGTCTTGTCTGTGGAGTG
TTGGACGAGCGTTCCCGCTGTGCCAGGTTGCCACCGGGAATGCGTCTGGGGCAGGAGGAGGACTCAGC
CGGTCTACAGGGCCAGACGCTAACAGGCGCCCGATCCGCATCGACTGGGAGGACTAGTTCGTAGCCATACTC
GAGTCCCTGCATTGGACGGCTCTGCTCTTTGGAGCTTCTCCCCCACCGCCCTTAAGAACATCTGCCAACAGC
TGGGTTACAGACTTCACTGTGAGTTCAGACTCCCAGCACCAACTCACTCTGATTCTGGTCCATTCACTGGGCA
CAGGTCACAGCACTGCTGAACAATGTGGCCTGGTGGGTTTCACTTTCTAGGTTGAAAACCTAACTGTCCA
CCCAGAAAGACACTCACCCATTCCCTCATTCTTTCTTACACTTAAATACCTCGTGTATGGTGAATCAGAC
CACAAAATCAGAAGCTGGGTATAATATTTCAAGTTACAAAACCTAGAAAAATTAACAGTTACTGAAATTATGA
CTTAAATACCCAATGACTCCTTAAATATGTAATATATAGTTATACCTTGAATTTCAATTCAAATGCAGACTAA
TTATAGGGAATTTGGAAGTGTATCAATAAACAGTATATAATTTT

FIGURE 110

MPPFLLLTCLFITGTSVSEFVALDPCSAYISLNEPWRNTDHQLDESQGPPLCDNHVNGEWYHFTGMAGDAMP
TFCIPENHCGTHAPVWLNGSHPLEGDGIVQRQACASFNNGCCLWNTTVEVKACPGGYVYRLTKPSVCFHV
YCGHFYDICEDECHGSCSDTSECTCAPGTVLGPDRQTCFDENECEQNNGGCSEICVNLKNSYRCECGVGRV
LRSDGKTCEDVEGCHNNNGGCSHSLGSEKGYQCECPRGLVLSEDNHTCQVPVLCKSNAIEVNI PRELVGG
LELFLTNTSCRGVSNATHVNI LFSKTCGTVVVDVNDKIVASNLVTGLPKQTPGSSGDFIIRTSKLLIPVT
CEFPRLYTISEGYVPNLRNSPLEIMSRNHGIFPFTLEIFKDNEFEPEPYREALPTLKLKRDLSLYFGIEPVVHV
SGLESLVESCFATPTS KIDEVLKYYLIRDGCVSDDSVKQYTSRDHLAKHFQVPVFKFVGGKHKEVFLHCRV
LVCVGLDERSRCAQGGCHRRMRGAGGEDSAGLQGGTTLTGGPIRIDWED

Important features of the protein:

Signal peptide:

amino acids 1-16

N-glycosylation sites.

amino acids 89-93, 116-120, 259-263, 291-295, 299-303

Tyrosine kinase phosphorylation sites.

amino acids 411-418, 443-451

N-myristoylation sites.

amino acids 226-232, 233-239, 240-246, 252-258, 296-302, 300-306,
522-528, 531-537

Aspartic acid and asparagine hydroxylation site.

amino acids 197-209

ZP domain proteins.

amino acids 431-457

Calcium-binding EGF-like proteins.

amino acids 191-212, 232-253

FIGURE 111

GAGAGAGGCAGCAGCTTGCTCAGCGGACAAGGATGCTGGGCGTGAGGGACCAAGGCCTGCCCTGCACTCGG
GCCTCCTCCAGCCAGTGTGACCAGGACTTCTGACCTGCTGGCCAGCCAGGACCTGTGTGGGGAGGCCCT
CCTGCTGCCTTGGGGTGACAATCTCAGCTCCAGGCTACAGGGAGACCGGGAGGATCACAGAGCCAGCATGT
TACAGGATCCTGACAGTGATCAACCTCTGAACAGCCTCGATGTCAAACCCTGCGCAAACCCCGTATCCCC
ATGGAGACCTTCAGAAAGGTGGGGATCCCCATCATCATAGCACTACTGAGCCTGGCGAGTATCATCATTGT
GGTTGTCTCATCAAGGTGATTCTGGATAAATACTACTTCTCTGCGGGCAGCCTCTCCACTTCATCCCGA
GGAAGCAGCTGTGTGACGGAGAGCTGGACTGTCCCTTGGGGGAGGACGAGGAGCACTGTGTCAAGAGCTTC
CCGAAGGGCCTGCAGTGGCAGTCCGCCCTCTCAAGGACCGATCCACACTGCAGGTGCTGGACTCGGCCAC
AGGGAAGTGGTTCTCTGCCTGTTTCGACAACCTCACAGAAGCTCTCGCTGAGACAGCCTGTAGGCAGATGG
GCTACAGCAGAGCTGTGGGATTGGCCAGACCAGGATCTGGATGTTGTTGAAATCACAGAAAACAGCCAG
GAGCTTCGCATGCGGAACCTCAAGTGGGCCCTGTCTCTCAGGCTCCCTGGTCTCCCTGCACTGTCTTGCCTG
TGGGAAGAGCCTGAAGACCCCCGTGTGGTGGGTGGGGAGGAGCCTCTGTGGATTCTTGGCCTTGGCAGG
TCAGCATCCAGTACGACAAACAGCACGTCTGTGGAGGGAGCATCCTGGACCCCACTGGGTCTCAGCGCA
GCCCACTGCTTCAGGAAACATAACCGATGTGTTCAACTGGAAGGTGCGGGCAGGCTCAGACAACTGGGCAG
CTTCCCATCCCTGGCTGTGGCCAAGATCATCATCATTGAATTCAACCCCATGTACCCCAAAGACAATGACA
TCGCCCTCATGAAGCTGCAGTTCCTCACTTCTCAGGCACAGTCAGGCCATCTGTCTGCCCTTCTTT
GATGAGGAGCTCACTCCAGCCACCCCACTCTGGATCATTGGATGGGCTTTACGAAGCAGAATGGAGGGAA
GATGTCTGACATACTGTGTCAGGCGTCAGTCCAGGTCATTGACAGCACACGGTGCAATGCAGACGATGCCGT
ACCAGGGGGAAGTCACCAGAGAAGATGATGTGTGCAGGCATCCCGGAAGGGGTGTGGACACCTGCCAGGGT
GACAGTGGTGGGCCCTGATGTACCAATCTGACCAGTGGCATGTGGTGGGCATCGTTAGCTGGGGCTATGG
CTGCGGGGGCCGAGCACCCAGGAGTATACCAAGGTCTCAGCCTATCTCAACTGGATCTACAATGTCT
GGAAGGCTGAGCTGTAATGCTGCTGCCCTTTGCAGTGTGGGAGCCGCTTCCTTCCTGCCCTGCCACCT
GGGGATCCCCAAAGTCAGACACAGAGCAAGAGTCCCCTTGGGTACACCCTCTGCCACAGCCTCAGCAT
TTCTTGGAGCAGCAAAGGGCCTCAATTCCTGTAAGAGACCCTCGCAGCCAGAGGGCCAGAGGAAGTCA
GCAGCCCTAGCTCGGCCACACTTGGTGTCTCCAGCATCCAGGAGAGACACAGCCCACTGAACAAGGTCT
CAGGGGTATTGCTAAGCCAAGAAGGAACTTTCCCACTACTGAATGGAAGCAGGCTGTCTTGTAAGGCC
CAGATCACTGTGGGCTGGAGAGGAGAAGGAAAGGGTCTGCGCCAGCCCTGTCCGTCTTACCCATCCCCAA
GCCTACTAGAGCAAGAAACAGTTGTAATATAAAAATGCACTGCCCTACTGTTGGTATGACTACCGTTACCT
ACTGTGTTCATTGTTATTACAGCTATGGCCACTATTATTAAAGAGCTGTGTAACATCTCTGGCAAAAAAA
AAAA

FIGURE 113

GGCTGGACTGGAACCTCTGGTCCCAAGTGATCCACCCGCCTCAGCCTCCCAAGGTGCTGTGATTA
TAGGTGTAAGCCACCGTGTCTGGCCTCTGAACAACCTTTTTCAGCAACTAAAAAGCCACAGGAGT
TGAAGTCTAGGATCTGACTATGCTGTTGGTGGCTAGTGTCTCTACTCTACCTACATTAATC
TGTTTTTTGTTCTCTGTAAGTACCTTTTACCTTCCTAACACAGAGGATCTGTCACTGTGGCTCT
GGCCAAACCTGACCTTCACTCTGGAACGAGAACAGAGGTTTCTACCCACACCGTCCCCTCGAAG
CCGGGGACAGCCTCACCTTGCTGGCCTCTCGCTGGAGCAGTGCCCTCACCAACTGTCTCACGTCT
GGAGGCACTGACTCGGGCAGTGCAGGTAGCTGAGCCTCTTGGTAGCTGCGGCTTTCAAGGTGGGC
CTTGCCCTGGCCGTAGAAGGGATTGACAAGCCGAAGATTTTCATAGGCGATGGCTCCCACTGCC
AGGCATCAGCCTTGCTGTAGTCAATCACTGCCCTGGGGCCAGGACGGGCCGTGGACACCTGTCTCA
GAAGCAGTGGGTGAGACATCACGCTGCCCGCCATCTAACCTTTTTCATGTCTGCACATCACCTG
ATCCATGGGCTAATCTGAAGTCTGTCCCAAGGAACCCAGAGCTTGAGTGAGCTGTGGCTCAGACC
CAGAAGGGGTCTGCTTAGACCACCTGGTTTTATGTGACAGGACTTGCAATCTCTGGAACATGAGG
GAACCCGGAGGAAAGCAAAGTGGCAGGGAAGGAACCTTGCCAAATATGGGTGAGAAAAGATG
GAGGTGTTGGGTATCACAAGGCATCGAGTCTCTGCATTCAGTGGACATGTGGGGAAAGGGCTG
CCGATGGCGCATGACACACTCGGGACTCACCTCTGGGGCCATCAGACAGCCGTTTCCGCCCGAT
CCACGTACCAGCTGCTGAAGGGCAACTGCAGGCCGATGCTCTCATCAGCCAGGCAGCAGCCAAAA
TCTGCGATCACCAGCCAGGGGCAGCCGTCTGGGAAGGAGCAAAGTGACCATTCTCCTCCC
CTCCTTCCCTCTGAGAGGCCCTCCTATGTCCCTACTAAAGCCACCAGCAAGACATAGCTGACAGG
GGCTAATGGCTCAGTGTGGCCAGGAGGTCAGCAAGGCCCTGAGAGCTGATCAGAAGGGCCTGCT
GTGCGAACACGGAAATGCCTCCAGTAAGCACAGGCTGCAAAATCCCAGGCAAAGGACTGTGTGG
CTCAATTTAAATCATGTTCTAGTAATTGGAGCTGTCCCAAGACCAAAGGAGCTAGAGCTTGTT
CAAATGATCTCCAAGGGCCCTTATACCCAGGAGACTTTGATTTGAATTTGAAACCCCAAATCCA
AACCTAAGAACCAGGTGCATTAAGAATCAGTTATTGCCGGGTGTGGTGGCCTGTAATGCCAACAT
TTTGGGAGGCCGAGGCGGTAGATCACCTGAGGTGAGGAGTTCAAGACCAGCCTGGCCAACATGG
TGAAACCCCTGTCTCTACTAAAAATACAAAAAACTAGCCAGGCATGGTGGTGTGTGCCTGTATC
CCAGCTACTCGGGAGGCTGAGACAGGAGAATTACTTGAACCTGGGAGGTGAAGGAGGCTGAGACA
GGAGAATCACTTCAGCCTGAGCAACACAGCGAGACTCTGTCTCAGAAAAATAAAAAAGAATTA
TGTTTATTTGTA

FIGURE 114

MLWWLVLLLLPTLKSVFCSLVTSLYLPNTEDLSLWLWPKPDLHSGTRTEVSTHTVPSKPGTASPC
WPLAGAVPSPPTVSRLEALTRAVQVAEPLGSCGFQGGPCPGRRRD

Signal peptide:
amino acids 1-15

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

FIGURE 115

CAGCAGTGGTCTCTCAGTCCTCTCAAAGCAAGGAAAGAGTACTGTGTGCTGAGAGACCATGGCAA
AGAATCCTCCAGAGAATTGTGAAGACTGTCACATTCTAAATGCAGAAGCTTTTAAATCCAAGAAA
ATATGTAAATCACTTAAGATTTGTGGACTGGTGTGGTATCCTGGCCCTAACTCTAATTGTCCT
GTTTGGGGGAGCAAGCACTTCTGGCCGGAGGTACCCAAAAAGCCTATGACATGGAGCACACTT
TCTACAGCAATGGAGAGAAGAAGATTACATGGAAATTGATCCTGTGACCAGAACTGAAATA
TTCAGAAGCGGAAATGGCACTGATGAAACATTGGAAGTGCACGACTTTAAAAACGGATACACTGG
CATCTACTTCGTGGGTCTTCAAAAATGTTTTATCAAACTCAGATTAAAGTGATTCTGAATTTT
CTGAACCAGAAGAGGAAATAGATGAGAATGAAGAAATTACCACAACCTTTCTTTGAACAGTCAGTG
ATTTGGGTCCCAGCAGAAAAGCCTATTGAAAACCGAGATTTTCTTAAAAATCCAAAATCTGGA
GATTTGTGATAACGTGACCATGTATTGGATCAATCCCACTCTAATATCAGTTTCTGAGTTACAAG
ACTTTGAGGAGGAGGAGAGATCTTCACTTTCTGCCAACGAAAAAAGGGATTGAACAAAAT
GAACAGTGGGTGGTCCCTCAAGTGAAGTAGAGAAGACCCGTCACGCCAGACAAGCAAGTGAGGA
AGAACTTCCAATAAATGACTATACTGAAAATGGAATAGAATTTGATCCCATGCTGGATGAGAGAG
GTTATTGTTGTATTTACTGCCGTCGAGGCAACCGCTATTGCCGCCGCTGTGTAACCTTTACTA
GGCTACTACCCATATCCATACTGCTACCAAGGAGGACGAGTCATCTGTGTCATCATGCCTTG
TAACTGGTGGTGGCCCGCATGCTGGGAGGGTCTAATAGGAGGTTTGGCTCAAATGCTTAAAC
TGCTGGCAACATATAATAAATGCATGCTATTCAATGAATTTCTGCCTATGAGGCATCTGGCCCT
GGTAGCCAGCTCTCCAGAATTACTTGTAGTAATTCCTCTCTCATGTTCTAATAAACTTCTACA
TTATCACCAAAAAAAAAAAAAAAAAA

FIGURE 116

MAKNPPENCEDCHILNAEAFKSKKICKSLKICGLVFGILALTLIVLFWGSKHFWPEVPPKAYDME
HTFYNSGEKKKIYMEIDPVTRTEIFRSGNGTDETLVHDFKNGYTGIFYVGLQKCFIKTQIKVIP
EFSEPEEEIDENEEITTTFFEQSVIWPVPAEKPIENRDFLKNKILEICDNVTMYWINPTLISVSE
LQDFEEEGEDLHFPANEKKGIEQNEQWVVPQVKVEKTRHARQASEEELPINDYTENGIEFDPMLD
ERGYCCICYCRRGNRYCRRVCEPLLGYYPYPYCYQGGRVICRVIMPCNWWVARMLGRV

Important features of the protein:

Signal peptide:

amino acids 1-40

Transmembrane domain:

amino acids 25-47 (type II)

N-glycosylation sites.

amino acids 94-97, 180-183

Glycosaminoglycan attachment sites.

amino acids 92-95, 70-73, 85-88, 133-136, 148-151, 192-195, 239-242

N-myristoylation sites.

amino acids 33-38, 95-100, 116-121, 215-220, 272-277

Microbodies C-terminal targeting signal.

amino acids 315-317

Cytochrome c family heme-binding site signature.

amino acids 9-14

FIGURE 117

GAGCTCCCCCAGGAGCGGTTAGCTTCACACCTTCGGCAGCAGGAGGGCGGCAGCTTCTCGCAGGGCGCA
GGCGGGCGGCCAGGATCATGTCACCACCACATGCCAAGTGGTGGCGTTCCTCCTGTCCATCCTGGGGCT
GGCCGGCTGCATCGCGGCCACCGGGATGGACATGTGGAGCACCAGGACCTGTACGACAACCCCGTCACCT
CCGTGTTCCAGTACGAAGGGCTCTGGAGGAGCTGCGTGAGGCAGAGTTCAGGCTTCACCGAATGCAGGCC
TATTTACCATCCTGGGACTTCCAGCCATGCTGCAGGCAGTGGAGCCCTGATGATCGTAGGCATCGTCT
GGGTGCCATTGGCCTCCTGGTATCCATCTTTGCCCTGAAATGCATCCGCATTGGCAGCATGGAGGACTCTG
CCAAAGCCAACATGACACTGACCTCCGGGATCATGTTTCATTGTCTCAGGCTTTGTGCAATTGCTGGAGTG
TCTGTGTTTGCCAACATGCTGGTGACTAACTTCTGGATGTCCACAGCTAACATGTACACCGGCATGGGTGG
GATGGTGACACTGTTCCAGACCAGGTACACATTTGGTGCGGCTCTGTTGCTGGGCTGGGTGCGCTGGAGGCC
TCACACTAATTGGGGTGTGATGATGTGCATCGCCTGCCGGGGCTGGCACCAGAAGAAACCACTACAAA
GCCGTTTCTTATCATGCCTCAGGCCACAGTGTGCCTACAAGCCTGGAGGCTTCAAGGCCAGCACTGGCTT
TGGGTCCAACACCAAAAACAAGAAGATATACGATGGAGGTGCCCGCACAGAGGACGAGGTACAATCTTATC
CTTCCAAGCAGCACTATGTGTAATGCTCTAAGACCTCTCAGCACGGGCGGAAGAACTCCCGGAGAGCTCA
CCAAAAACAAGGAGATCCCATCTAGATTTCTTCTGCTTTTACTCACAGCTGGAAGTTAGAAAAGCCT
CGATTTATCTTTGGAGAGGCCAAATGGTCTTAGCCTCAGTCTCTGTCTCTAAATATTTCCACCATAAAACA
GCTGAGTTATTTATGAATTAGAGGCTATAGCTCACATTTTCAATCCTCTATTTCTTTTTTAAATATAACT
TTCTACTCTGATGAGAGAATGTGGTTTTAATCTCTCTCACATTTTGATGATTTAGACAGACTCCCCCTC
TTCTCCTAGTCAATAAACCCATTGATGATCTATTTCCAGCTTATCCCAAGAAAACCTTTGAAAGGAAA
GAGTAGACCCAAAGATGTTATTTCTGCTGTTGAATTTGTCTCCCCACCCCAACTGGCTAGTAATAA
ACACTTACTGAAGAAGAAGCAATAAGAGAAAGATATTTGTAATCTCTCCAGCCATGATCTCGGTTTTCTT
ACACTGTGATCTTAAAAGTTACCAAACCAAGTCATTTTCAGTTTGAGGCAACCAAACCTTTCTACTGCTG
TTGACATCTTCTTATTACAGCAACACCATTCTAGGAGTTTCTGAGCTCTCCACTGGAGTCCTCTTCTGT
CGCGGTCAGAAATGTCCCTAGATGAATGAGAAAATTTTTTTTTTAATTAAGTCCATAATATAGTTAA
AATAAATAATTTTTAGTAAAATGATACACTATCTCTGTGAAATAGCCTCACCCCTACATGTGGATAGAAG
GAAATGAAAAATAATTGCTTTGACATTTGCTATATGGTACTTTGTAAGTCATGCTTAAGTACAAATCC
ATGAAAAGCTCACACCTGTAATCCTAGCACTTTGGGAGGCTGAGGAGGAAGGATCACTTGAGCCAGAAAGT
TCGAGACTAGCCTGGGCAACATGGAGAAGCCCTGTCTCTACAAAATACAGAGAGAAAAAATCAGCCAGTCA
TGGTGGCATAACCTGTAGTCCCAGCATTCGGGAGGCTGAGGTGGGAGGATCACTTGAGCCAGGGAGGT
TGGGGCTGCAGTGAGCCATGATCACACCACGCACCTCCAGCCAGGTGACATAGCGAGATCCTGTCTAAAAA
AATAAAAAATAAATAATGGAACACAGCAAGTCTTAGGAAGTAGGTTAAAATAATCTTTAA

FIGURE 118

MSTTTCQVVAFLLSILGLAGCIAATGMDMWSTQDLYDNPVTSVVFQYEGLRSCVRQSSGFTECRP
YFTILGLPAMLQAVRALMIVGIVLGAIGLLVSI FALKCIRIGSMEDSAKANMTLTSGIMFIVSGL
CAIAGVSVFANMLVTNFWMSTANMYTGMGMVQTVQTRYTFGAALFVGWVAGGLTLIGGVMMCIA
CRGLAPEETNYKAVSYHASGHSVAYKPGGFKASTGFGSNTKNKKIYDGGARTEDEVQSYPSKHDY
V

Signal peptide:

amino acids 1-23

Transmembrane domains:

amino acids 81-100, 121-141, 173-194

FIGURE 119

GGAAAACTGTTCTCTTCTGTGGCACAGAGAACCCTGCTTCAAAGCAGAAGTAGCAGTTCCGGAGTCC
AGCTGGCTAAAACATCCCAGAGGATAATGGCAACCCATGCCTTAGAAATCGCTGGGCTGTTTCTTG
GTGGTGTGGAATGGTGGGCACAGTGGCTGTCACTGTCATGCCTCAGTGGAGAGTGTCCGGCCTTCATT
GAAAACAACATCGTGGTTTTTAAAACTTCTGGGAAGGACTGTGGATGAATTGCCGTGAGGCAGGCTAA
CATCAGGATGCAGTGCAAAATCTATGATTCCCTGTGGCTCTTCTCCGGACCTACAGGCAGCCAGAG
GACTGATGTGTGCTGCTTCCGTGATGTCCTTCTTGGCTTTCATGATGGCCATCCTTGGCATGAAATGC
ACCAGGTGCACGGGGACAATGAGAAGGTGAAGGCTCACATTCTGCTGACGGCTGGAATCATCTTCAT
CATCACGGGCATGGTGGTGTCTATCCCTGTGAGCTGGGTTGCCAATGCCATCATCAGAGATTTCTATA
ACTCAATAGTGAATGTTGCCAAAAACGTGAGCTTGAGAGAAGCTCTCTACTTAGGATGGACCACGGCA
CTGGTGTGATTGTTGGAGGAGCTCTGTCTGCTGCGTTTTTTGTTGCAACGAAAAGAGCAGTAGCTA
CAGATACTCGATACCTCCCATCGCACAAACCAAAAAAGTTATCACACCGGAAAGAAGTCACCGAGCG
TCTACTCCAGAAGTCAATGTAGTTGTGTATGTTTTTAACTTTACTATAAAGCCATGCAAATG
ACAAAAATCTATATTACTTTCTCAAATGGACCCAAAGAACTTTGATTTACTGTTCTTAACTGCCT
AATCTTAATTACAGGAACTGTGCATCAGCTATTTATGATTCTATAAGCTATTTCCAGCAGAATGAGATA
TTAAACCCAAATGCTTTGATTGTTCTAGAAAGTATAGTAATTTGTTTTCTAAGGTGGTTCAAGCATCTA
CTCTTTTTATCATTACTTCAAATGACATTGCTAAAGACTGCATTATTTACTACTGTAATTTCTCC
ACGACATAGCATTATGTACATAGATGAGTGTAAACATTTATATCTCACATAGAGACATGCTTATATGGT
TTTATTTAAATGAAATGCCAGTCCATTACACTGAATAAATAGAACTCAACTATTGCTTTTCAGGGAA
ATCATGGATAGGGTTGAAGAAGGTTACTATTAATGTTTTAAAAACAGCTTAGGGATTAATGTCCCTCCA
TTTATAATGAAGATTAATGAAGGCTTTAATCAGCATTGTAAAGGAAATTGAATGGCTTTCTGATAT
GCTGTTTTTTAGCCTAGGAGTTAGAAATCCTAACTTCTTTATCCTCTTCTCCAGAGGCTTTTTTTTT
CTTGTGTATTAAATTAACATTTTTAAAACGCAGATATTTGTCAAGGGGCTTTGCATTCAACTGCTT
TTCCAGGGCTATACTCAGAAGAAAGATAAAAGTGTGATCTAAGAAAAAGTATGTTTTTAGGAAAGTG
AAAAATTTTTGTTTTGTTTGAAGAAGAATGATGCATTTTGACAAGAAATCATATATGTATGGAT
ATATTTAATAAGTATTTGAGTACAGACTTTGAGGTTTCATCAATATAAATAAAAGAGCAGAAAAATA
TGTCTTGGTTTTTCATTTGCTTACCAAAAAACAACAACAAAAAAGTTGTCCTTTGAGAACCTCACCT
GCTCCTATGTGGGTACCTGAGTCAAATGTCAATTTTTGTTCTGTGAAAAATAAATTTCCCTCTTGTA
CCATTTCTGTTTAGTTTTACTAAAATCTGTAATACTGTATTTTTCTGTTTATCCAAATTTGATGAA
ACTGACAATCCAATTTGAAAGTTTGTGTCGACGCTGTCTAGCTTAAATGAATGTGTTCTATTTGCTT
TATACATTTATATTAATAAATGTACATTTTTCTAATT

FIGURE 120

MATHALEIAGLFLGGVGMVGTVAVTVMPQWRVSAFIENNVVVFENFWEGLWMNCVRQANIRMQCK
IYDSSLALSPDLQAARGLMCAASVMSFLAFMMAILGMKCTRCTGDNEKVKAHILLTAGIIFIITG
MVLVLPVSWVANAIIRDYNSIVNVAQKRELGEALYLGWTTALVLI VGGALFCCVFCCNEKSSSY
RYSIPSHRTTQKSYHTGKKSPSVYSRSQYV

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 82-101, 118-145, 164-188

FIGURE 121

GGAGAGAGGCGCGGGTCAAAGGCGCATTGATGCAGCCTGCGGGCGCCTCGGAGCGCGGGAG
CCAGACGCTGACCACGTTCTCTCCTCGGTCTCCTCCGCTCCAGCTCCGCGCTGCCCGGCAGCC
GGGAGCCATGCACCCCAGGGCCCCGCCCTCCCCGAGCGGCTCCGCGGCCTCTGCTGCTCC
TGCTGCTGCAGCTGCCCGCGCCGTGAGCGCCTCTGAGATCCCCAAGGGGAAGCAAAGGCGCAG
CTCCGGCAGAGGGAGGTGGTGGACCTGTATAATGGAATGTGCTTACAAGGGCCAGCAGGAGTGCC
TGGTCGAGACGGGAGCCCCTGGGGCCAATGTTATTCCGGGTACACCTGGGATCCCAGGTCCGGATG
GATTCAAAGGAGAAAAGGGGAATGTCTGAGGAAAGCTTTGAGGAGTCTGGACACCCAACTAC
AAGCAGTGTTCATGGAGTTCATTGAATTATGGCATAGATCTTGGGAAAATTGGGAGTGTACATT
TACAAAGATGCGTTCAAATAGTGCTCTAAGAGTTTTGTTTCAAGTGGCTCACTTCGGCTAAAATGCA
GAAATGCATGCTGTCAGCGTTGGTATTTACATTCAATGGAGCTGAATGTTTCAAGACCTCTTCCC
ATTGAAGCTATAATTTATTTGGACCAAGGAAGCCCTGAAATGAATCAACAATTAATATTCATCG
CACTTCTTCTGTGGAAGGACTTTGTGAAGGAATTGGTGCTGGATTAGTGGATGTTGCTATCTGGG
TTGGCACTTGTTCAGATTACCCAAAAGGAGATGCTTCTACTGGATGGAATTCAGTTTCTCGCATC
ATTATTGAAGAACTACCAAAATAAATGCTTTAATTTTCAATTTGCTACCTCTTTTTTTATTATGCC
TTGGAATGGTTCACTTAAATGACATTTTAAATAAGTTTATGTATACATCTGAATGAAAAGCAAAG
CTAAATATGTTTACAGACCAAAGTGTGATTTACACTGTTTTTAAATCTAGCATTATTCATTTTG
CTTCAATCAAAGTGGTTTCAATATTTTTTTTAGTTGGTTAGAATACTTTCTTCATAGTCACATT
CTCTCAACCTATAATTTGGAATATTGTTGTGGTCTTTTGTTTTTTCTCTTAGTATAGCATTTTTA
AAAAAATATAAAAGCTACCAATCTTTGTACAATTTGTAATGTTAAGAATTTTTTTTATATCTGT
TAAATAAAAATTATTTCCAACA

FIGURE 122

MRPQGPAASPQRLRGLLLLLLLLQLPAPSSASEIPKGGKQKAQLRQREVVDLYNGMCLQGPAGVPGR
DGSPGANVIPGTPGIPGRDGFKEKGECLRESFEESWTPNYKQCSWSSLNYGIDLKIAECTFTK
MRSNSALRVLFSGSLRLKCRNACCQRWYFTFNGAECGGLPIEAIYLDQGSPEMNSTINIHRYS
SVEGLCEGIGAGLVDVAIWVGTCSDYPKGDASTGWNSVSRIIIIEELPK

Signal peptide:

amino acids 1-30

Transmembrane domain:

amino acids 195-217

FIGURE 123

GCTGAGCGTGTGCGCGGTACGGGGCTCTCCTGCCTTCTGGGCTCCAACGCAGCTCTGTGGCTGAA
CTGGGTGCTCATCACGGAACTGCTGGGCTATGGAATACAGATGTGGCAGCTCAGGTAGCCCCAA
ATTGCTGGAAGAATACATCATGTTTTTCGATAAGAAGAAATTGTAGGATCCAGTTTTTTTTTTA
ACCGCCCCCTCCCACCCCCCAAAAAACTGTAAAGATGCAAAAACGTAATATCCATGAAGATCC
TATTACCTAGGAAGATTTTGATGTTTTGCTGCGAATGCGGTGTTGGGATTTATTTGTTCTTGGAG
TGTTCTGCGTGGCTGGCAAAGAATAATGTTCCAAAATCGGTCCATCTCCAAGGGGTCCAATTTT
TCTTCTGGGTGTGAGCGGCCCTGACTCACTACAGTGCAGCTGACAGGGGCTCATGCAACTG
GCCCCAAGCCAAAGCAAAGACCTAAGGACGACCTTTGAACAATACAAAGGATGGGTTTCAATG
TAATTAGGCTACTGAGCGGATCAGCTGTAGCACTGGTTATAGCCCCACTGTCTTACTGACAATG
CTTTCTTCTGCGAACGAGGATGCCCTAAGGGCTGTAGGTGTGAAGGCAAATGGTATATTGTGA
ATCTCAGAAATACAGGAGATACCCTCAAGTATATCTGCTGGTTGCTTAGGTTTGTCCCTTCGCT
ATAACAGCCTTCAAAAACCTAAGTATAATCAATTTAAAGGGCTCAACCAGCTCAGCTGGCTATAC
CTTGACCATAACCATATCAGCAATATTGACGAAAATGCTTTAATGGAATACGCAGACTCAAAGA
GCTGATTCCTTAGTTCCAATAGAATCTCTATTTTCTTAACAATACCTTCAGACCTGTGACAAAT
TACGGAACCTGGATCTGTCTATAATCAGCTGCATTCCTGGGATCTGAACAGTTTCGGGGCTTG
CGGAAGCTGCTGAGTTTACATTTACGGTCTAACTCCCTGAGAACCATCCCTGTGCGAATATTTCCA
AGACTGCCGCAACCTGGAACCTTTGGACCTGGGATATAACCGGATCCGAAGTTTAGCCAGGAATG
TCTTTGCTGGCATGATCAGACTCAAAGAACTTCACCTGGAGCACAAATCAATTTTCCAAGCTCAAC
CTGGCCCTTTTCCAAGTTGGTCAAGCTTTCAGAACCTTTACTTGCAAGTGAATAAAAATCAGTGT
CATAGGACAGACCATGTCTTGGACCTGGAGCTCCTTACAAAGGCTTGATTTATCAGGCAATGAGA
TCGAAGCTTTTCAGTGGACCCAGTGTTCCTCAGTGTGTCGCGAATCTGCAGCGCCTCAACCTGGAT
TCCAACAAGCTCACATTTATGGTCAAGAGATTTTGGATTCTTGGATATCCCTCAATGACATCAG
TCTTGTGGGAATATATGGGAATGCAGCAGAAATATTTGCTCCCTGTAAACTGGCTGAAAAGTT
TTAAAGGTCTAAGGGAGAATACAATTATCTGTGCCAGTCCCAAAGAGCTGCAAGGAGTAAATGTG
ATCGATGCAGTGAAGAATACAGCATCTGTGGCAAAGTACTACAGAGAGGTTTGATCTGGCCAG
GGCTCTCCCAAAGCCGACGTTAAGCCCAAGCTCCCCAGGCCGAAGCATGAGAGCAAACCCCTT
TGCCCCGACGGTGGGAGCCACAGAGCCCGGCCAGAGACCGATGCTGACGCCGAGCACATCTCT
TTCCATAAAATCATCGCGGGCAGCGTGGCGCTTTTCTGTCCGTGCTCGTCATCTGCTGGTTAT
CTACGTGTCATGGAAGCGGTACCCTGCGAGCATGAAGCAGCTGCAGCAGCGCTCCCTCATGCGAA
GGCACAGGAAAAGAAAAGACAGTCCCTAAAGCAAATGACTCCCAGCACCCAGGAATTTTATGTA
GATTATAAACCCACCAACACGGAGACCAGCGAGATGCTGTGAATGGGACGGGACCCCTGCACCTA
TAACAAATCGGGCTCCAGGGAGTGTGAGGTA**TGA**ACCATTGTGATAAAAAGAGCTCTTAAAGCT
GGGAAATAAGTGGTGCTTTATTGAACTCTGGTACTATCAAGGAAACGCGATGCCCCCTCCCC
TTCCCTCTCCCTCTCACTTTGGTGGCAAGATCCTTCTTGTCCGTTTTAGTGCATTATAATACT
GGTCATTTTCTCTCATAATAATCAACCCATTGAAATTTAAATACCACAATCAATGTGAAGCTT
GAACTCCGGTTAATATAATACCTATTGTATAAGACCCTTTACTGATTCCATTAATGTGCGATTT
GTTTTAAGATAAAACTTCTTTCATAGGTAAAAA

FIGURE 124

MGFNVIRLLSGSAVALVIAPTPLLTLSSAERGCPKGCRCCEGKMVYCESQKLQEI PSSISAGCLG
LSLRYNLSLQKLYNQFKGLNQLTWLYLDHNHISNIDENAFNGIRRLKELILSSNRISYFLNNTFR
PVTNLRNLDSYNQLHSLGSEQFRGLRKLSSLHLRSNSLRTPVRI FQDCRNLELLDLGYNRIRS
LARNVFAGMIRLKEHLHLEHNQFSKLNLFRLVSLQNLVQWNKISVIGQTMSTWSSLQRLDL
SGNEIEAFSGSPVFCVPLQRLNLDNKLTFIQEILDSWISLNDISLAGNIWECSRNICSLVN
WLKSFKGLRENTIICASPKEIQGVNVIDAVKNYSICGKSTTERFDLARALPKPTFKPKLPRPKHE
SKPPLPPTVGATEPGPETDADAEHISFHKI IAGSVALFSLVLLVILVIYVSWKRYPASMKQLQQR
SLMRRHRKKRQSLKQMTPTSTQEFYVDYKPTNTETSEMLLNGTGPCTYNKSGSRECEV

Important features of the protein:

Signal peptide:

amino acids 1-33

Transmembrane domain:

amino acids 420-442

N-glycosylation sites.

amino acids 126-129, 357-360, 496-499, 504-507

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 465-468

Tyrosine kinase phosphorylation site.

amino acids 136-142

N-myristoylation sites.

amino acids 11-16, 33-38, 245-250, 332-337, 497-502, 507-512

FIGURE 125

CCGTTATCGTCTTGCCTACTGCTGAATGTCCGTCCCAGGAGGAGGAGAGAGGCTTTTGCCGCTG
ACCCAGAGATGGCCCCGAGCGAGCAAATTCCTACTGTCCGGCTGCGCGGCTACCGTGGCCGAGCT
AGCAACCTTTCCCTGGATCTCACAAAACCTCGACTCCAATGCAAGGAGAAGCAGCTCTTGCTC
GGTTGGGAGACGGTGCAAGAGAATCTGCCCCCTATAGGGGAATGGTGCACAGCCCTAGGGATC
ATTGAAGAGGAAGGCTTTCTAAAGCTTTGGCAAGGAGTGACACCCGCCATTTACAGACACGTAGT
GTATTCTGGAGGTCGAATGGTCACATATGAACATCTCCGAGAGGTTGTGTTTGGCAAAAGTGAAG
ATGAGCATTATCCCCTTTGAAATCAGTCATTGGAGGGATGATGGCTGGTGTATTGGCCAGTTT
TTAGCCAATCCAACCTGACCTAGTGAAGGTTGAGATGCAAATGGAAGGAAAAGGAAACTGGAAGG
AAAACCATTGCGATTTCTGGTGTACATCATGCATTTGCAAAAATCTTAGCTGAAGGAGGAATAC
GAGGGCTTTGGGAGGCTGGGTACCCAATATACAAAGAGCAGCAGCTGGTGAATATGGGAGATTTA
ACCCTTATGATACAGTGAACACTACTTGGTATTGAATACACCACTTGAGGACAATATCATGAC
TCACGGTTTATCAAGTTTATGTTCTGGACTGGTAGCTTCTATTCTGGGAACACCAGCCGATGTCA
TCAAAAGCAGAATAATGAATCAACCACGAGATAAACAAGGAAGGGGACTTTTGTATAAATCATCG
ACTGACTGCTTGATTGAGGCTGTTCAAGGTTGAAGGATTCATGAGTCTATATAAAGGCTTTTTACC
ATCTTGGCTGAGAATGACCCCTTGGTCAATGGTGTCTGGCTTACTTATGAAAAAATCAGAGAGA
TGAGTGGAGTCAGTCCATTTTAA

FIGURE 126

MSVPEEEERLLPLTQRWPRASKFLLSGCAATVAELATFPLDLTKTRLQMQGEAALARLGDGARES
APYRGMVRTALGIIIEEGFLKLWQGVTPAIYRHVVYSGGRMVTYEHLREVVFVGKSEDEHYPLWKS
VIGGMMAGVIGQFLANPTDLVKVQMOMEKGRKLEKPLRFRGVHHAFAKILAEGGIRGLWAGWVP
NIQRAALVNMGLTTYDVTVKHYLVLNTPLEDNIMTHGLSSLCSGLVASILGTPADVIKSRIMNQP
RDKQGRGLLYKSSTDCLIQAVQGEFMSLYKGFPSWLRMTPWSMVFWLTYEKIREMSGVSPF

Transmembrane domains:

amino acids 25-38, 130-147, 233-248

FIGURE 127

CGCGGATCGGACCCAAGCAGGTCGGCGGGCGGGCAGGAGAGCGGCCGGGCGTCAGCTCCTCGAC
CCCCGTGTCGGGCTAGTCCAGCGAGGGCGGACGGGGCGGCGTGGGCCCATGGCCAGGCCCGGCATGG
AGCGGTGGCGCGACCGGCTGGCGCTGGTGACGGGGGCTCGGGGGGCATCGGCGCGGCCGTGGCC
CGGGCCCTGGTCCAGCAGGGACTGAAGGTGGTGGGCTGCGCCCGCACTGTGGGCAACATCGAGGA
GCTGGCTGCTGAATGTAAGAGTGCAGGCTACCCCGGGACTTTGATCCCTACAGATGTGACCTAT
CAAATGAAGAGGACATCCTCTCCATGTTCTCAGCTATCCGTTCTCAGCACAGCGGTGTAGACATC
TGCATCAACAATGCTGGCTTGGCCCGGCTGACACCCTGCTCTCAGGCAGCACCAGTGGTTGGAA
GGACATGTTCAATGTGAACGTGCTGGCCCTCAGCATCTGCACACGGGAAGCCTACCAGTCCATGA
AGGAGCGGAATGTGGACGATGGGCACATCAATAACATCAATAGCATGTCTGGCCACCGAGTGTTA
CCCCTGTCTGTGACCCACTTCTATAGTGCCACCAAGTATGCCGTCACTGCGCTGACAGAGGGACT
GAGGCAAGAGCTTCGGGAGGCCAGACCCACATCCGAGCCACGTGCATCTCTCCAGGTGTGGTGG
AGACACAATTCGCCTTCAAACCTCCACGACAAGACCCTGAGAAGGCAGCTGCCACCTATGAGCAA
ATGAAGTGTCTCAAACCCGAGGATGTGGCCGAGGCTGTATCTACGTCTCAGCACCCCGCACA
CATCCAGATTGGAGACATCCAGATGAGGCCACGGAGCAGGTGACCTAGTGACTGTGGGAGCTCC
TCCTTCCCTCCCCACCCTTCATGGCTTGCCCTCTGCCTCTGGATTTTAGGTGTTGATTTCTGGAT
CACGGGATACCACTTCCGTCCACACCCCGACCAGGGGCTAGAAAATTTGTTTGAGATTTTATA
TCATCTTGTCAAATTGCTTCAGTTGTAATGTGAAAAATGGGCTGGGAAAGGAGGTGGTGTCCC
TAATTGTTTTACTTGTTAACTTGTCTTGTGCCCTGGGCACTTGGCCTTTGTCTGCTCTCAGTG
TCTTCCCTTTGACATGGGAAAGGAGTTGTGGCCAAAATCCCCATCTTCTTGACCTCAACGCTG
TGCTCAGGGCTGGGGTGGCAGAGGGAGGCCTTACCTTATATCTGTGTTGTTATCCAGGGCTCC
AGACTTCCTCCTCTGCCTGCCCCACTGCACCCTCTCCCCTTATCTATCTCCTTCTCGGCTCCCC
AGCCCAGTCTTGCTTCTGTCCCCTCCTGGGGTCATCCCTCCACTCTGACTCTGACTATGGCAG
CAGAACCAGGGCCTGGCCAGTGGATTTTCATGGTGATCATTAAAAAGAAAAATCGCAACCAA
AAAAAAAAA

FIGURE 128

MARPGMERWRDRDLALVTGASGGIGAAVARALVQQGLKVVGCARTVGNIEELAAECKSAGYPGTLI
PYRCDLSNEEDIILSMFSAIRSQHSVDICINNAGLARPDTLISGSGTSGWKDMFNVNLALSICTR
EAYQSMKERNVDDGHIININSMGHRVLPVSVTHFYATKYAVTALTEGLRQELREAQTHIRATC
ISPGVVETQFAFKLHDKDPEKAAATYEQMKCLKPEDVAEAVIYVLSIPAHIQIGDIQMRPTEQVT

Important features of the protein:

Signal peptide:

amino acids 1-17

N-myristoylation sites.

amino acids 18-24, 21-27, 22-28, 24-30, 40-46, 90-96, 109-115,
199-205

Short-chain alcohol dehydrogenase.

amino acids 30-42, 104-114

FIGURE 129

AACTTCTACATGGGCCTCCTGCTGCTGGTGTCTTCCCTCAGCCTCCTGCCGGTGGCCTACACCAT
CATGTCCCTCCCACCCTCCTTTGACTGCGGGCCGTTCCAGTGCAGAGTCTCAGTTGCCCGGGAGC
ACCTCCCCTCCCAGGCAGTCTGCTCAGAGGGCCTCGGCCAGAATCCAGTTCGGTTTCATGC
CAGCCTGTAAAAGGCCATGGAACTTTGGGTGAATCACCGATGCCATTAAAGAGGGTTTTCTGCCA
GGATGGAAATGTTAGGTCGTTCTGTGTCTGCGCTGTTCAATTCAGTAGCCACCAGCCACCTGTGG
CCGTTGAGTGCTTGAAATGAGGAACTGAGAAAATTAATTTCTCATGTATTTTTCTCATTATTTA
TTAATTTTTAACTGATAGTTGTACATATTTGGGGTACATGTGATATTTGGATACATGTATACAA
TATATAATGATCAAATCAGGGTAACTGGGATATCCATCACATCAAACATTTATTTTTTATTCTTT
TTAGACAGAGTCTCACTCTGTCACCCAGGCTGGAGTGCAGTGGTGCCATCTCAGCTTACTGCAAC
CTCTGCCTGCCAGGTTCAAGCGATTCTCATGCCTCCACCTCCCAAGTAGCTGGGACTACAGGCAT
GCACCACAATGCCCACTAATTTTTGTATTTTTAGTAGAGACGGGGTTTTGCCATGTTGCCCAGG
CTGGCCTTGAACTCCTGGCCTCAAACAATCCACTTGCCTCGGCCCTCCAAAGTGTATGATTACA
GGCGTGAGCCACCGTGCCTGGCCTAAACATTTATCTTTCTTTGTGTGGGAACTTTGAAATTAT
ACAATGAATTATTGTTAACTGTATCTCCCTGCTGTGCTATGGAACACTGGGACTTCTTCCCTCT
ATCTAACTGTATATTTGTACCAGTTAACCAACCGTACTTCATCCCCACTCCTCTCTATCTTCCC
AACCTCTGATCACCTCATTCTACTCTCTACCTCCATGAGATCCACTTTTTTAGCTCCCACATGTG
AGTAAGAAAATGCAATATTTGTCTTTCTGTGCCTGGCTTATTTCACTTAACATAATGACTTCCTG
TTCCATCCATGTTGCTGCAAATGACAGGATTTTCGTTCTTAATTTCAATTAATAAATAACCACACATG
GCAAAAA

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

FIGURE 130

MGLLLLVLFLSLLPVAYTIMSLPPSFDCGPFRCRVSVAREHLPSRGSLLRGPRPRIPVLVSCQPV
KGGHTLGESPMPPFKRVFCQDGNVRSFCVCAVHFSSHQPPVAVECLK

Important features of the protein:

Signal peptide:

amino acids 1-18

N-myristoylation site.

amino acids 86-92

Zinc carboxypeptidases, zinc-binding region 2 signature.

amino acids 68-79

FIGURE 131

TTCTGAAGTAACGGAAGCTACCTTGTATAAAGACCTCAACACTGCTGACCATGATCAGCGCAGCCTGGAGC
ATCTTCCTCATCGGGACTAAAATTGGGCTGTTCCCTCAAGTAGCACCTCTATCAGTTATGGCTAAATCCTG
TCCATCTGTGTGTCGCTGCGATGCGGGTTTCATTTACTGTAATGATCGCTTCTGACATCCATTCCAACAG
GAATACCAGAGGATGCTACAACCTCTACCTTCAGAACAACCAAATAAATAATGCTGGGATTCCCTCAGAT
TTGAAAACTTGCTGAAAGTAGAAAGAATATACCTATACCACAACAGTTTAGATGAATTTCTACCAACCT
CCCAAAGTATGTAAGAGTTACATTTGCAAGAAAATAACATAAGGACTATCACTTATGATTCACTTTCAA
AAATTCCTATCTGGAAGAATTACATTTAGATGACAACCTCTGTCTCTGCAGTTAGCATAGAAGAGGGAGCA
TTCCGAGACAGCAACTATCTCCGACTGCTTTTCCCTGTCCCGTAATCACCTTAGCACAATTCCTGGGGTTT
GCCCAGGACTATAGAAGAACTACGCTTGATGATAATCGCATATCCACTATTTTCATCACCATCTCTCAAG
GTCTCACTAGTCTAAAACGCCTGGTTCAGATGGAACCTGTTGAACAATCATGGTTTAGGTGACAAAGTT
TTCTTCAACCTAGTTAATTTGACAGAGCTGTCCCTGGTGGGAATTCCTGACTGCTGCACCAGTAAACCT
TCCAGGCACAAACCTGAGGAAGCTTTATCTTCAAGATAACCACATCAATCGGGTGCCCCAAATGCTTTTT
CTTATCTAAGGCAGCTCTATCGACTGGATATGTCCAATAATAACCTAAGTAATTTACCTCAGGGTATCTTT
GATGATTTGGACAATATAACACAACCTGATTTCTCGCAACAATCCCTGGTATTGCGGGTGCAAGATGAAATG
GGTACGTGACTGGTTACAATCACTACCTGTGAAGGTCAACGTGCGTGGGCTCATGTGCCAAGCCCCAGAAA
AGGTTCTGTTGGGATGGCTATTAAGGATCTCAATGCAGAACCTGTTTGATTGTAAGGACAGTGGGATGTAAGC
ACCATTGAGATAACCACTGCAATACCCAACACAGTGTATCCTGCCAAGGACAGTGGCCAGCTCCAGTGAC
CAAACAGCCAGATATTAAGAACCCCAAGCTCACTAAGGATCAACAAACCACAGGGAGTCCCTCAAGAAAA
CAATTACAATTACTGTGAAGTCTGTCACTCTGATACCATTTCATATCTCTTGAAAACCTTGCTCTACCTATG
ACTGCTTTGAGACTCAGCTGGCTTAAACTGGGCCATAGCCCGCATTTGGATCTATAACAGAAACAATTGT
AACAGGGGAACGCAGTGAAGTACTTGGTCACAGCCCTGGAGCCTGATTCACCTATAAAGTATGCATGGTTC
CCATGGAACCAGCAACCTCTACCTATTTGATGAAACTCCTGTTTGTATTGAGACTGAAACTGCACCCCTT
CGAATGTACAACCCACACCCTCAATCGAGAGCAAGAGAAAGAACCTTACAAAACCCCAATTTACC
TTTGGCTGCCATCATTTGGTGGGGCTGTGGCCCTGGTTACCATTGCCCTTCTTGCTTTAGTGTGTTGGTATG
TTCATAGGAATGGATCGCTCTTCTCAAGGAACCTGTGCATATAGCAAAGGGAGGAGAAGAAAGGATGACTAT
GCAGAAGCTGGCACTAAGAAGGACAACCTCTATCCTGGAATCAGGGAACTTCTTTTTCAGATGTTACCAAT
AAGCAATGAACCCATCTCGAAGGAGGAGTTTGTAAATACACACCATATTTCCCTCCTAATGGAATGAATCTGT
ACAAAAACAATCACAGTGAAGCAGTAGTAACCGAAGCTACAGAGACAGTGGTATTCCAGACTCAGATCAC
TCACACTCATGATGCTGAAGGACTCACAGCAGACTTGTGTTTTGGGTTTTTAAACCTAAGGGAGGTGATG
GT

FIGURE 132

MISAAWSIFLIGTKIGLFLQVAPLSVMAKSCPSVCRCDAGFIYCNDRELTISIPTGIPEDATTLYL
QNNQINNAGIPSDLKLLKVERIYLYHNSLDEFPTNLPKYVKELHLQENNIRTIYDLSKIPYL
EELHLDNSVSAVSIIEGAFRDSNYLRLLEFLSRNHLSTIPWGLPRTIEELRLDDNRISTISSPSL
QGLTSLKRLVLDGNLLNHHGLGDKVFFNLVNLTELSLVRNSLTAAPVNLPGTNLRKLYLQDNHIN
RVPPNAFSYLRQLYRLDMSNNLSNLPQGI FDDLDNITQLILRNNPWYCGCKMKWVRDWLQSLPV
KVNVRGLMCQAPEKVRGMAIKDLNAELFDCKDSGIVSTIQITTAIPNTVYPAQGQWPAPVTKQPD
IKNPKLTKDQQTGSPSRKTIITITVKSVTSDTIHISWKLALPMTALRLSWLKLGHSPAFGSITET
IVTGERSEYLVTALEPDSPYKVMVPMETSNLVLFDETPVCIETETAPLRMYNPTTTLNREQEKE
PKYKPNLPLAAIIGGAVALVTIALLALVCWYVHRNGSLFSRNCAYSKGRRRKDDYAEAGTKKDNS
ILEIRETSFQMLPISNEPISKEEFVIHTIFPPNGMNLKNNHSESSSNRSYRDSGIPDSHSHS

Important features of the protein:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 531-552

N-glycosylation sites.

amino acids 226-229, 282-285, 296-299, 555-558, 626-629, 633-636

Tyrosine kinase phosphorylation site.

amino acids 515-522

N-myristoylation sites.

amino acids 12-17, 172-177, 208-213, 359-364, 534-539, 556-561,
640-645

Amidation site.

amino acids 567-570

Leucine zipper pattern.

amino acids 159-180

Phospholipase A2 aspartic acid active site.

amino acids 34-44

FIGURE 133

CCGTCATCCCCCTGCAGCCACCCTTCCCAGAGTCCTTTGCCAGGCCACCCAGGCTTCTTGCCA
GCCCTGCCGGGCCACTTGTCTTCATCTCTGCCAGGGGAGGTGGGAAGGAGGTGGGAGGAGGGCG
TGCAGAGGCAGTCTGGGCTTGGCCAGAGCTCAGGGTGTGAGCGTGTGACCAGCAGTGAGCAGAG
GCCGGCCATGGCCAGCCTGGGGCTGTCTCTCTACTGACAGCACTGCCACCGCTGTGGT
CCTCCTCACTGCCTGGGCTGGACACTGTGAAAGTAAAGCCACCATTGCAGACCTGATCCTGTCT
GCGCTGGAGAGAGCCACCGTCTTCCCTAGAACAGAGGCTGCCTGAAATCAACCTGGATGGCATGGT
GGGGGTCCGAGTGTGGAAGAGCAGCTAAAAAGTGTCCGGGAGAAGTGGGCCAGGAGCCCTGC
TGCAGCCGCTGAGCCTGCGCGTGGGGATGCTGGGGGAGAAGCTGGAGGCTGCCATCCAGAGATCC
CTCCACTACCTCAAGCTGAGTGATCCCAAGTACCTAAGAGAGTCCAGCTGACCCTCCAGCCCGG
GTTTTGGAAGCTCCCACATGCCTGGATCCACACTGATGCCTCCTTGGTGTACCCACGTTCCGGGC
CCCAGGACTCATTCTCAGAGGAGAGAAGTGACGTGTGCCTGGTGCAGCTGTGGAAACCGGGACG
GACAGCAGCGAGCCCTGCGGCCTCTCAGACCTCTGCAGGAGCCTCATGACCAAGCCCGGCTGCTC
AGGCTACTGCCTGTCCCACCAACTGCTCTTCTTCTCTGGGCCAGAATGAGGGGATGCACACAGG
GACCACTCCAACAGAGCCAGGACTATATCAACCTCTTCTGCGCCAACATGATGGACTTGAACCGC
AGAGCTGAGGCCATCGGATACGCTACCTACCCGGGACATCTTCATGGAAAACATCATGTTCTG
TGAATGGGCGGCTTCTCCGACTTCTACAAGCTCCGGTGGCTGGAGGCCATTCTCAGCTGGCAGA
AACAGCAGGAAGGATGCTTCGGGGAGCCTGATGCTGAAGATGAAGAATTATCTAAAGCTATTCAA
TATCAGCAGCATTTCGAGGAGAGTGAAGAGGCGAGAAAACAATTTCCAGATTCTCGCTCTGT
TGCTCAGGCTGGAGTACAGTGGCGCAATCTCGGCTCACTGCAACCTTTGCCTCCTGGGTTCAAGC
AATTCTCTTGCCTCATCTCCCAGTAGCTGGGACTACAGGAGCGTGCCACCATACCTGGCTAAT
TTTTATATTTTTTAGTAGAGACAGGGTTTCATCATGTTGCTCATGCTGGTCTCGAACTCCTGAT
CTCAAGAGATCCGCCACCTCAGGCTCCCAAAGTGTGGGATTATAGGTGTGAGCCACCGTGTCTG
GCTGAAAAGCACTTTCAAAAGAGACTGTGTTGAATAAAGGGCCAAGGTTCTTGCCACCCAGCACTC
ATGGGGGCTCTCTCCCTAGATGGCTGCTCTCCACAACACAGCCACAGCAGTGGCAGCCCTGG
GTGGCTTCCATACATCTTGGCAGAATACCCCCAGCAAACAGAGAGCCACACCCATCCACACCG
CCACCACCAAGCAGCCGCTGAGACGGACGGTTCATGCCAGCTGCCTGGAGGAGGAACAGACCCC
TTTAGTCTCATCCCTTAGATCCTGGAGGGCACGGATCACATCCTGGGAAGAAGGCATCTGGAGG
ATAAGCAAAGCCACCCGACACCCAATCTTGAAGCCCTGAGTAGGCAGGGCCAGGGTAGGTGGG
GGCCGGGAGGGACCCAGGTGTGAACGGATGAATAAAGTTCAACTGCAACTGAAAAA

FIGURE 134

MSARGRWEGGRRACRGSGLLARAQGAERVTSSSEQRPMASLGLLLLLLLLTALPPLWSSSLPGLD
TAESKATIADLILSALERATVFLEQRLPEINLDGMVGVVLEEQLKSVREKWAQEPPLQPLSLRV
GMLGEKLEAAIQRSLHYLKLSDPKYLREFQLTLQPGFWKLPHAWIHTDASLVYPTFGPQDSFSEE
RSDVCLVQLLGTGTDSSSEPCGLSDLCRSLMTPGCSGYCLSHQLLFFLWARMRGCTQGPLQSQD
YINLFCANMMDLNRRAEAIQYAYPTRDI FMENIMFCMGGFSDFYKLRWLEAILSWQKQEGCFG
EPDAEDEELSKAIQYQQHFSRRVKREKQFPDSRSVAQAGVQWRNLGSLQPLPPGFKQFSCLLILP
SSWDYRSVPPYLANFYIFLVETGFHHVAHAGLELLISRDPTSGSQSVGL

Important features of the protein:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 39-56

Tyrosine kinase phosphorylation sites.

amino acids 149-156, 274-282

N-myristoylation sites.

amino acids 10-16, 20-26, 63-69, 208-214

Amidation site.

amino acids 10-14

Glycoprotein hormones beta chain signature 1.

amino acids 230-237

FIGURE 135

GGTCTGAGTGCAGAGCTGCTGTCAATGGCGGCCGCTCTGTGGGGCTTCTTTCCCGTCCTGCTGCTG
CTGCTGCTATCGGGGGATGTCCAGAGCTCGGAGGTGCCCGGGGCTGCTGCTGAGGGATCGGGAGG
GAGTGGGGTCGGCATAGGAGATCGCTTCAAGATTGAGGGGCGTGCAGTTGTTCCAGGGGTGAAGC
CTCAGGACTGGATCTCGGCGGCCCGAGTGCTGGTAGACGGAGAAGAGCACGTGCGTTTCTTAAG
ACAGATGGGAGTTTTGTGGTTCATGATATACCTTCTGGATCTTATGTAGTGAAGTTGTATCTCC
AGCTTACAGATTTGATCCCGTTCGAGTGGATATCACTTCGAAAGGAAAAATGAGAGCAAGATATG
TGAATTACATCAAACATCAGAGGTTGTCAGACTGCCCTATCCTCTCAAATGAAATCTTCAGGT
CCACCTTCTTACTTTATTAAAAGGGAATCGTGGGGCTGGACAGACTTTCTAATGAACCCAATGGT
TATGATGATGGTTCTTCCTTTATTGATATTTGTGCTTCTGCCTAAAGTGGTCAACACAAGTGATC
CTGACATGAGACGGGAAATGGAGCAGTCAATGAATATGCTGAATFCCAACCATGAGTTGCCTGAT
GTTTCTGAGTTCATGACAAGACTCTTCTCTTCAAATCATCTGGCAAATCTAGCAGCGGCAGCAG
TAAAACAGGCAAAAGTGGGGCTGGCAAAGGAGGTAGTAGCCGTCAGAGCTGGCATTGTCAC
AAACACGGCAACACTGGGTGGCATCCAAGTCTTGAAAACCGTGTGAAGCAACTACTATAAACTT
GAGTCATCCCGACGTTGATCTCTTACAACCTGTGTATGTT
AACTTTTTAGCACATGTTTTGTACTTGGTACACGAGAAAACCCAGCTTTCATCTTTTGTCTGTAT
GAGGTCAATATTGATGTCACTGAATTAATTACAGTGTCTATAGAAAATGCCATTAATAAATAT
ATGAACTACTATACATTATGTATATTAATTAACATCTTAATCCAGAAATCAAAAAAAAAAAAA
AAAAAAAAAAAAAA

FIGURE 136

MAAALWGFFPVLLLLLLSGDVQSSEVPGAAAEGSGGSGVGIGDRFKIEGRAVVPGVKPDWISAA
RVLVDGEEHVGFLKTDGSEFVVHDIPSGSYVVEVVSAPYRFDVPRVDITSKGMRARYVNYIKTSE
VVRLPYPLOMKSSGPPSYFIKRESWGWTDFLMNPMVMMMLVPLLI FVLLPKVVNTSDPDMRREME
QSMNMLNSNHELDPDVSEFMTRLFSSKSSGKSSSGSSKTKGSGAGKRR

Important features of the protein:

Signal sequence:

amino acids 1-23

Transmembrane domain:

amino acids 161-182

N-glycosylation site.

amino acids 184-187

Glycosaminoglycan attachment sites.

amino acids 37-40, 236-239

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 151-154

N-myristoylation sites.

amino acids 33-38, 36-41, 38-44, 229-234

Amidation site.

amino acids 238-241

ATP/GTP-binding site motif A (P-loop).

amino acids 229-236

FIGURE 137

GATGGCGCAGCCACAGCTTCTGTGAGATTCGATTTCTCCCAGTTCCTGTGGGTCTGAGGGGA
CCAGAAGGGTGAGCTACGTTGGCTTTCTGGAAGGGGAGGCTATAATGCGTCAATCCCCAAAACAA
GTTTTGACATTTCCCCTGAAATGTCATTCTCTATCTATTCAGTCAAGTGCCTGTGTTCCAGGC
CTTACCTGCTGGGCACTAACGGCGGAGCCAGGATGGGGACAGAATAAAGGAGCCACGACCTGTGC
CACCAACTCGCACTCAGACTCTGAACTCAGACCTGAAATCTTCTCTCACGGGAGGCTTGGCAGT
TTTTCTTACTCCTGTGGTCTCCAGATTTTCAGGCCTAAGATGAAAGCCTCTAGTCTTGCCCTCAGC
CTTCTCTGTGCGTTTTATCTCCTATGGACTCCTTCCACTGGACTGAAGACACTCAATTTGGG
AAGCTGTGTGATCGCCACAAACCTTCAGGAAATACGAAATGGATTTTCTGAGATACGGGGCAGTG
TGCAAGCCAAAGATGAAACATTGACATCAGAATCTTAAGGAGGACTGAGTCTTTGCAAGACACA
AAGCCTGCGAATCGATGCTGCCCTCGCCCATTTGCTAAGACTCTATCTGGACAGGGTATTTAA
AACTACCAGACCCCTGACCATTATACTCTCCGGAAGATCAGCAGCCTCGCCAATTCCTTTCTTA
CCATCAAGAAGGACCTCCGGCTCTCTCATGCCACATGACATGCCATTGTGGGGAGGAAGCAATG
AAGAAATACAGCCAGATTCTGAGTCACTTTGAAAAGCTGGAACCTCAGGCAGCAGTTGTGAAGGC
TTTGGGGGAAC TAGACATTTCTGCAATGGATGGAGGAGACAGAATAGGAGGAAAGTGATGCTG
CTGCTAAGAATATTCGAGGTCAAGAGCTCCAGTCTTCAATACCTGCAGAGGAGGCATGACCCCAA
ACCACCATCTCTTACTGTACTAGTCTTGTGCTGGTCACAGTGTATCTTATTTATGCATTACTTG
CTTCCTTGCAATGATGCTTTATGCATCCCCAATCTTAATTGAGACCATACTTGTATAAGATTTT
TGTAATATCTTCTGCTATTGGATATATTTATTAGTTAATATATTTATTTATTTTTGCTATTTA
ATGTATTTATTTTTTACTTGGACATGAACTTTAAAAAAATTCACAGATTATATTATAACCTG
ACTAGAGCAGGTGATGTATTTTTATACAGTAAAAAAAACCTTGTAATCTAGAAGAGTGG
CTAGGGGGTTATTCATTTGTATTCAACTAAGGACATATTTACTCATGCTGATGCTCTGTGAGAT
ATTTGAAATTGAACCAATGACTACTTAGGATGGTGTGGAATAAGTTTTGATGTGGAATTGCAC
ATCTACCTTACAATTACTGACCATCCCCAGTAGACTCCCCAGTCCATAATTGTGTATCTTCCAG
CCAGGAATCTACACGGCCAGCATGTATTTCTACAAATAAAGTTTTCTTTGCATACCAAAAAAAA
AAAAAAAAAA

FIGURE 138

MRQFPKTSFDISPEMSFSIYSLQVPAVPG LTCWALTAEPGWGQNGATTTCATNSHSDSELRPEIF
SSREAWQFFLLWSPDFRPMKASSLAFSLLSAAFYLLWTPSTGLKTLNLGSCVIATNLQEIRNG
FSEIRGSVQAKDGNIDIRILRRTESLQDTK PANRCCLLRHLLRLYLDRVFKNYQTPDHYYTLRKIS
SLANSFLTIIKKDLRLSHAHMTCHCGEEAMKKYSQILSHFEKLEPQAAVVKALGELDILLOWMEET
E

Important features of the protein:

Signal peptide:

amino acids 1-42

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 192-195, 225-228

N-myristoylation sites.

amino acids 42-47, 46-51, 136-141

FIGURE 140

MRLGSGTFATCCVAIEVLGIAVFLRGFFPAPVRSSARAEGAEPPAPEPSAGASSNWTTLPPPLF
SKVVIVLIDALRDDFVFGSKGVKFMPTTYLVEKGASHSFVAEAKPPTVTMPRIKALMTGSLPGF
VDVIRNLNSPALLEDVIRQAKAAGKRIVFYGDETWWKLFPKHFVEYDGTTSFFVSDYTEVDNNV
TRHLDKVLKRGDWDILILHYLGLDGHIGHISGPN SPLIGQKLSEMDSVLMKIHTSLQSKERETPLP
NLLVLCGDHGMSETGSHGASSTEEVNTPLILISSAFERKPGDIRHPKHVQ

Important features of the protein:

Signal peptide:

amino acids 1-34

Transmembrane domain:

amino acids 58-76

N-glycosylation sites.

amino acids 56-60, 194-198

N-myristoylation sites.

amino acids 6-12, 52-58, 100-106, 125-131, 233-239, 270-276,
275-281, 278-284

Amidation site.

amino acids 154-158

Cell attachment sequence.

amino acids 205-208

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

FIGURE 142

MLLLLLLEYNFPIENNCQHLKTHTFRVKNLNPKKFSIHDQDHKVLVLD SGNLI AVDPK NYIRPEI
FFALASSLSSASAÆKGSPI LLGVSKGEFCLYCDKDKGQSHPSLQLKKEKLMKLA AQKESARRPFI
FYRAQVGSWNMLESAHPGWFICTSCNCNEPVGVTDKFENRKHIEFSFQPVCKAEMSPSEVSD

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 33-36

N-myristoylation site.
amino acids 50-55, 87-92

Interleukin-1
amino acids 37-182

FIGURE 144

MLGLPWKGGLSWALLLLLLLGSQILLIYAWHFHEQRDCDEHNVMARYLPATVEFAVHTFNQOSKDY
YAYRLGHILNSWKEQVESKTVFSMELLGRTRCGKFEDDIDNCHFQESTELNNTFTCCFFTISTRP
WMTQFSLLNKTCLEGFH

Important features of the protein:

Signal peptide:

amino acids 1-25

N-glycosylation sites.

amino acids 117-121, 139-143

N-myristoylation site.

amino acids 9-15

FIGURE 145

CTGTGCAGCTCGAGGCTCCAGAGGCACACTCCAGAGAGAGCCAAGGTTCTGACGCGATGAGGAAG
CACCTGAGCTGGTGGTGGCTGGCCACTGTCTGCATGCTGCTCTTCAGCCACCTCTCTGCGGTCCA
GACGAGGGGCATCAAGCACAGAATCAAGTGAACCGGAAGGCCCTGCCAGCACTGCCAGATCA
CTGAGGCCAGGTGGCTGAGAACC GCCGGGAGCCTTCATCAAGCAAGGCCGCAAGCTCGACATT
GACTTCGGAGCCGAGGGCAACAGGTACTACGAGGCCAACTACTGGCAGTTCCCCGATGGCATCCA
CTACAACGGCTGCTCTGAGGCTAATGTGACCAAGGAGGCATTTGTACCAGGCTGCATCAATGCCA
CCCAGGCCGCGAACCAGGGGAGTTCCAGAAGCCAGACAACAAGCTCCACCAGCAGGTGCTCTGG
CGGCTGGTCCAGGAGCTCTGCTCCCTCAAGCATTGCGAGTTTTGGTTGGAGAGGGGCGCAGGACT
TCGGGTACCATGCACCAGCCAGTGCTCCTCTGCCCTTCTGGCTTTGATCTGGCTCATGGTGAAT
AAGCTTGCCAGGAGGCTGGCAGTACAGAGCGCAGCAGCGAGCAAATCCTGGCAAGTGACCCAGCT
CTTCTCCCCAAACCCACGCGTGTCTGAAGGTGCCAGGAGCGGCGATGCACTCGCACTGCAAA
TGCCGCTCCCACGTATGCGCCCTGGTATGTGCCTGCGTTCTGATAGATGGGGACTGTGGCTTCT
CCGTCCTCCATTCTCAGCCCCTAGCAGAGCGTCTGGCACACTAGATTAGTAGTAAATGCTTGAT
GAGAAGAACACATCAGGCACTGCGCCACCTGCTTACAGTACTTCCAACAACCTTTAGAGGTAG
GTGTATTCCCGTTTTACAGATAAGGAACTGAGGCCAGAGAGCTGAAGTACTGCACCCAGCATC
ACCAGCTAGAAAGTGGCAGAGCCAGGATCAACCCTGGCTTGTCTAACCCAGGTTTTCTGCTCT
GTCCAATCCAGAGCTGTCTGGTGATCACTTTATGTCTCACAGGGACCCACATCCAACATGTAT
CTCTAATGAAATTGTGAAAGCTCCATGTTTAAATAAATGAAAACACCTGA

FIGURE 146

MRKHLSSWWLATVCMLLFSHLSAVQTRGIKHKRIRKWNRKALPSTAQITEAQVAENRPGAFIKQGRK
LDIDFGAEGNRYYEANYWQFPDGIHYNGCSEANVTKEAFVTGCINATQAANQGEFQKPDNKLHQQ
VLWRLVQELCSLKHCFWLERGAGLRVTMHQPVLCLLALLIWLMMVK

Important features of the protein:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 157-171

N-glycosylation sites.

amino acids 98-102, 110-114

Tyrosine kinase phosphorylation site.

amino acids 76-83

N-myristoylation sites.

amino acids 71-77, 88-94, 93-99, 107-113, 154-160

Amidation site.

amino acids 62-66

FIGURE 147

GCCTTGGCCTCCCAAAGGGCTGGGATTATAGGCGTGACCACCATGTCTGGTCCAGAGTCTCATTT
CCTGATGATTTATAGACTCAAAGAAAACTCATGTTCAGAAGCTCTCTTCTCTCTGGCCTCCTCT
CTGTCTTCTTTCCCTCTTCTTCTTATTTAATTAGTAGCATCTACTCAGAGTCATGCAAGCTGG
AAATCTTTCATTTTGCTTGTGAGTGGGGTAGGTCACTGAGTCTTAGTTTTTATTTTTGAAATTT
CAACTTTCAGATTCAGGGGGTACATGTGAAGGTTTGTTTTATGAGTATATTGCATTGATGCTGAGG
TTTGGGGT

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

FIGURE 148

MFRSSLLFWPPLCLLSLFLILLISSIYSESKLEIFHFACQWGRSLSLSFYFLKFQLSDSGGTCE
GLFYEYIA

Important features of the protein:

Signal peptide:

amino acids 1-25

N-myristoylation site.

amino acids 62-68

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

FIGURE 149

GTCTCCGCGTCACAGGAACCTTCAGCACCCACAGGGCGGACAGCGCTCCCCTCTACCTGGAGACTTGAC
TCCCGCGCGCCCCAACCTGCTTATCCCTTGACCGTGCAGAGATCCTGCAGCCGCCAGTCC
CGGCCCTCTCCCGCCCCACACCACCCCTCTGGCTCTTCTGTCTTTACTCCTCCTTTTCATTTCATA
ACAAAAGCTACAGCTCCAGGAGCCAGCGCGGGCTGTGACCCAAGCCGAGCGTGAAGAATGGGGTT
CCTCGGGACCGGCACTTGATTCTGGTGTAGTGCTCCCGATTCAAGCTTTCCCCAAACCTGGAGGAA
GCCAAGACAATCTCTACATAATAGAGAATTAAGTGCAGAAAGACCTTTGAATGAACAGATTGCTGAA
GCAGAAGAAGACAAGATTAAAAACATATCCTCCAGAAAACAAGCCAGGTGAGAGCAACTATTCTTT
TGTTGATAACTTGAACCTGCTAAAGGCAATAACAGAAAAGGAAAAAATGAGAAAGAAAGACAATCTA
TAAGAAGCTCCCCTTGATAATAAGTGAATGTGAAGATGTTGATTCAACCAAGAATCGAAAAGT
ATCGATGATTATGACTCTACTAAGAGTGGATTGGATCATAAATTTCAAGATGATCCAGATGGTCTTCA
TCAACTAGACGGGACTCCTTTAACCCTGAAGACATGTCCATAAAATCGCTGCCAGGATTTATGAAG
AAAATGACAGAGCCGTGTTTGACAAGATTGTTTCTAAACTACTTAATCTCGGCCTTATCACAGAAAGC
CAAGCACATACACTGGAAGATGAAGTAGCAGAGGTTTACAAAAATTAATCTCAAAGGAAGCCAACAA
TTATGAGGAGGATCCCAATAAGCCACAAGCTGGACTGAGAATCAGGCTGGAAAAATACCAGAGAAAG
TGACTCCAATGGCAGCAATTCAAGATGGTCTTGCTAAGGGAGAAAACGATGAAACAGTATCTAACACA
TTAACCTTGACAAATGGCTTGAAAAGGAGAACTAAAACCTACAGTGAAGACAACCTTTGAGGAACTCCA
ATATTTCCCAAATTTCTATGCGCTACTGAAAAGTATTGATTCAGAAAAAGAAGCAAAGAGAAAGAAA
CACTGATTACTATCATGAAAACACTGATTGACTTTGTGAAGATGATGGTGAATATGGAACAATATCT
CCAGAAGAAGGTGTTTCTACCTTGAAAACCTGGATGAAATGATTGCTCTTCAGACCAAAAACAAGCT
AGAAAAAATGCTACTGACAATATAAGCAAGCTTTTCCAGCACCATCAGAGAAGAGTCATGAAGAAA
CAGACAGTACCAAGGAAGAAGCAGCTAAGATGGAAAAGGAATATGGAAGCTTGAAGGATTCACAAAA
GATGATAACTCCAACCAGGAGGAAAGACAGATGAACCCAAAGGAAAAACAGAAGCCTATTTGGAAGC
CATCAGAAAAAATATTGAATGGTTGAAGAAACATGACAAAAAGGAAATAAAGAAGATTATGACCTTT
CAAAGATGAGAGACTTCATCAATAAACAAGCTGATGCTTATGTGGAGAAAGGCATCCTTGACAAGGAA
GAAGCCGAGGCCATCAAGCGCATTATAGCAGCCTGTAAAAATGGCAAAGATCCAGGAGTCTTTCAA
CTGTTTCAGAAAACATAATATAGCTTAAAACACTTCTAATCTGTGATTAAAAATTTTTTGACCAAGG
GTTATTAGAAAGTCTGAATTTACAGTAGTTAACCTTTTACAAGTGGTTAAAACATAGCTTTCTCCC
GTAAAAACTATCTGAAAGTAAAGTTGTATGTAAGCTGAAAAAAAAAAAAAAAAAAAAA

FIGURE 150

MGFLGTGTWILVVLVLP IQAFPKPGGSQDKSLHNRELSAERPLNEQIAEAEEDKIKKTYPPENKPG
QSNYSFVDNLNLLKAI TEKEKIEKERQSIRSSPLDNKLNVEDVDSTKNRKLIDDYDSTKSGLDHK
FQDDPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVFDKIVSKLLNLGLITESQAHTLEDEVAE
VLQKLISKEANNYEEDPNKPTSWTENQAGKIPEKVTPMAAIQDGLAKGENDETVSNTLTLTNGLE
RRTKTYSEDNFEELQYFPNFYALLKSIDSEKEAKEKETLITIMKTLIDFVKMMVKYGTISPEEGV
SYLENLDEMIALQTKNKLEKNATDNISKLFPAPEKSHEETDSTKEEA AKMEKEYGSLKDSTKDD
NSNPGGKTDEPKGKTEAYLEAIRKNIEWLKKHKGNKEDYDLSKMRDFINKQADAYVEKGILDK
EEAEAIKRIYSSL

N-glycosylation sites:

amino acids 68-71, 346-349, 350-353

Casein kinase II phosphorylation site:

amino acids 70-73, 82-85, 97-100, 125-128, 147-150, 188-191, 217-
220, 265-268, 289-292, 305-308, 320-323, 326-329, 362-365, 368-
341, 369-372, 382-385, 386-389, 387-390

N-myristoylation sites:

amino acids 143-148, 239-244

FIGURE 151

CGGCTCGAGGCTCCCGCCAGGAGAAAGGAACATTCTGAGGGGAGTCTACACCCGTGGAGCTCAA
GATGGTCTCTGAGTGGGGCGCTGTGCTTCCGAATGAAGGACTCGGCATTGAAGGTGCTTTATCTGC
ATAATAACCAGCTTCTAGCTGGAGGGCTGCATGCAGGGAAGGTCAATTAAGGTGAAGAGATCAGC
GTGGTCCCCAATCGGTGGCTGGATGCCAGCCTGTCCCCGTATCCTGGGTGTCCAGGGTGGAAAG
CCAGTGCCTGTCATGTGGGGTGGGGCAGGAGCCGACTCTAACACTAGAGCCAGTGAACATCATGG
AGCTCTATCTTGGTGCCAAGGAATCCAAGAGCTTACCTTCTACCGGCGGGACATGGGGCTCACC
TCCAGCTTCGAGTCGGCTGCCTACCCGGGCTGGTTCCCTGTGCACGGTGCCTGAAGCCGATCAGCC
TGTCAGACTCACCCAGCTTCCCGAGAATGGTGGCTGGAAATGCCCCATCACAGACTTCTACTTCC
AGCAGTGTGACTAGGGCAACGTGCCCCCAGAACTCCCTGGGCAGAGCCAGCTCGGGTGGGGGT
GAGTGGAGGAGACCCATGGCGGACAATCACTCTCTCTGCTCTCAGGACCCCCACGCTGACTTAG
TGGGCACCTGACCACTTTGTCTTCTGGTTCCCGATTGGATAAAATCTGAGATTTGGAGCTCAGT
CCACGGTCTCCCCACTGGATGGTGTACTGCTGTGGAACCTTGTAAAAACCATGTGGGGTAAA
CTGGGAATAACATGAAAAGATTTCTGTGGGGTGGGGTGGGGGAGTGGTGGGAATCATCTCTGCT
TAATGGTAACTGACAAGTGTACCTGAGCCCCGAGCCCAACCCATCCCCAGTTGAGCCTTATA
GGGTGAGTAGTCTCCACATGAAGTCTGTCACCTACCAGTGTGCAGGAGAGGGAGGTGGTCAATA
GAGTCAGGGATCTATGGCCCTTGGCCAGCCCCACCCCTTCCCTTAAATCCTGCCACTGTCAATA
TGCTACCTTTCCTATCTCTTCCCTCATCATCTTGTGTGGGCATGAGGAGGTGGTGTATGTCAGAA
GAAATGGCTCGAGCTCAGAAGATAAAAAGATAAGTAGGGTATGCTGATCCTCTTTAAAAACCCAA
GATACAATCAAAATCCCAGATGCTGGTCTCTATTCCCATGAAAAGTGTCTATGACATATTGAGA
AGACCTACTTACAAAGTGGCATATATTGCAATTTATTTAATTAAGATAACCTATTTATATATT
TCTTTATAGAAAAAGTCTGGAAGAGTTACTTCAATTGTAGCAATGTGAGGGTGGTGGCAGTAT
AGGTGATTTTTCTTTTAAATCTGTTAATTTATCTGTATTTCCCTAATTTTTCTACAATGAAGATGA
ATTCTTGTATAAAAAATAAGAAAAGAAATTAATCTTGAGGTAAGCAGAGCAGACATCATCTCTGA
TTGTCTCAGCCTCCACTTCCCCAGAGTAAATTCAAATTGAATCGAGCTCTGCTGCTCTGGTTGG
TTGTAGTAGTGATCAGGAAACAGATCTCAGCAAAGCCACTGAGGAGGAGGCTGTGCTGAGTTTGT
GTGGCTGGAATCTCTGGGTAAGGAACCTAAAGAACAAAAATCATCTGGTAATCTTTCTTAGAAG
GATCACAGCCCCCTGGGATTCCAAGGCATGGATCCAGTCTCTAAGAAGGCTGCTGTACTGGTTGA
ATTGTGTCCCCCTCAAATTCACATCCTTCTTGGAAATCTCAGTCTGTGAGTTTATTGGAGATAAG
GTCTCTGCAGATGTAGTTAGTTAAGACAAGGTCTGCTGGATGAAGGTAGACCTAAATTCATAT
GACTGGTTTTCTTGTATGAAAAGGAGAGGACACAGAGACAGAGGAGACGCGGGGAGACTATGTA
AAGATGAAGGCAGAGATCGGAGTTTTGCAGCCACAAGCTAAGAAACACCAAGGATGTGGCAACC
ATCAGAAGCTTGAAGAGGCAAAGAAGATTCTTCCCTAGAGGCTTTAGAGGGATAACGGCTCTG
CTGAAACCTTAATCTCAGACTTCCAGCCTCCTGAACGAAGAAGAATAAATTCGGCTGTTTTAA
GCCACCAAGGATAATTGGTTACAGCAGCTCTAGGAACTAATACAGCTGCTAAAAATGATCCCTGT
CTCCTCGTGTTTACATCTGTGTGTGTCCCCTCCCACAATGTACCAAAGTTGTCTTTGTGACCAA
TAGAATATGGCAGAAGTATGGCATGCCACTTCCAAGATTAGGTTATAAAAGACTGCAGCTTC
TACTTGAGCCTCTCTCTGCCACCCACCGCCCCAATCTATCTGGCTCACTCGCTCTGGGGG
AAGCTAGCTGCCATGCTATGAGCAGGCCTATAAAGAGACTTACGTGGTAAAAATGAAGTCTCCT
GCCACAGCCACATTAGTGAACCTAGAAGCAGAGACTCTGTGAGATAATCGATGTTTGTGTTTT
AAGTTGCTCAGTTTTGGTCTAACTTGTATGCAGCAATAGATAAAATAATATGCAGAGAAAGAG

FIGURE 153

CTTCAGAACAGGTTCTCCTTCCCCAGTCACCAGTTGCTCGAGTTAGAATTGTCTGCAATGGCCGC
CCTGCAGAAATCTGTGAGCTCTTTCCTTATGGGGACCCTGGCCACCAGCTGCCTCCTTCTCTTGG
CCCTCTTGGTACAGGGAGGAGCAGCTGCGCCCATCAGCTCCCAGTGCAGGCTTGACAAGTCCAAC
TTCCAGCAGCCCTATATACCAACCGCACCTTCATGCTGGCTAAGGAGGCTAGCTTGGCTGATAA
CAACACAGACGTTTCGTCTCATTGGGGAGAACTGTTCCACGGAGTCAGTATGAGTGAGCGCTGCT
ATCTGATGAAGCAGGTGCTGAACTTACCCTTGAAGAAGTGCTGTCCCTCAATCTGATAGGTTT
CAGCCTTATATGCAGGAGGTGGTGGCCCTTCTGGCCAGGCTCAGCAACAGGCTAAGCACATGTCA
TATTGAAGGTGATGACCTGCATATCCAGAGGAATGTGCAAAGCTGAAGGACACAGTAAAAAGC
TTGGAGAGAGTGGAGAGATCAAAGCAATTGGAGAACTGGATTGCTGTTTATGTCTCTGAGAAAT
GCCTGCATTGACCAGAGCAAAGCTGAAAAATGAATAACTAACCCCTTTCCTGCTAGAAATAA
CAATTAGATGCCCCAAAGCGATTTTTTTTAAACAAAAGGAAGATGGGAAGCCAAACTCCATCATG
ATGGGTGGATTCCAAATGAACCCCTGCGTTAGTTACAAAGGAAACCAATGCCACTTTTGTTTATA
AGACCAGAAGGTAGACTTCTAAGCATAGATATTTATTGATAACATTTTATTGTAAGTGGTGTTC
TATACACAGAAAACAATTTATTTTTTAAATAATTGCTTTTTTCCATAAAAAAGATTACTTTCCAT
TCCTTTAGGGGAAAAAACCCCTAAATAGCTTCATGTTTCCATAATCAGTACTTTATATTTATAAA
TGTATTTATTATTATTATAAGACTGCATTTTATTATATCATTTTATTAATATGGATTTATTTAT
AGAAACATCATTCGATATTGCTACTTGAGTGTAAGGCTAATATTGATATTTATGACAATAATTAT
AGAGCTATAACATGTTTATTTGACCTCAATAAACACTTGGATATCCC

FIGURE 154

MAALQKSVSSFLMGTLATSCLLLLLALLVQGGAAAPISSHCRLDKSNFQOPYITNRTFMLAKEASL
ADNNTDVRLLIGEKLFGVSMSERCYLMKQVLNFTLEEVLPQSDRFQPYMQEVVPFLARLSNRLS
TCHIEGDDLHIQRNVQKLKDTVKKLGESGEIKAIGELDLLFMSLRNACI

Important features of the protein:

Signal peptide:

amino acids 1-33

N-glycosylation sites.

amino acids 54-58, 68-72, 97-101

N-myristoylation sites.

amino acids 14-20, 82-88

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 10-21

FIGURE 155

GGCTTGCTGAAAATAAAATCAGGACTCCTAACCTGCTCCAGTCAGCCTGCTTCCACGAGGCCTGT
CAGTCAGTGCCCGACTTGTGACTGAGTGTGCAGTGCCAGCATGTACCAGGTGAGTGCAGAGGGC
TGCTGAGGGCTGTGCTGAGAGGGAGAGGAGCAGAGATGCTGCTGAGGGTGGAGGGAGGCCAAGC
TGCCAGGTTTGGGGCTGGGGGCCAAGTGGAGTGAAGAACTGGGATCCCAGGGGGAGGGTGCAGAT
GAGGGAGCGACCCAGATTAGGTGAGGACAGTTCTCTCATTAGCCTTTTCCCTACAGGTGGTTGCAT
TCTTGGCAATGGTTCATGGGAACCCACACTACAGCCACTGGCCCAGCTGCTGCCCCAGCAAAGGG
CAGGACACCTCTGAGGAGCTGCTGAGGTGGAGCACTGTGCCTGTGCCTCCCCTAGAGCCTGCTAG
GCCCAACCGCCACCCAGAGTCCCTGTAGGGCCAGTGAAGATGGACCCCTCAACAGCAGGGCCATCT
CCCCCTGGAGATATGAGTTGGACAGAGACTTGAACCGGCTCCCCAGGACCTGTACCACGCCCCGT
TGCCTGTGCCCCGACTGCGTCAGCCTACAGACAGGCTCCACATGGACCCCGGGCAACTCGGA
GCTGCTCTACCACAACCAGACTGTCTTCTACAGGCGGCCATGCCATGGCGAGAAGGGCACCCACA
AGGGCTACTGCCTGGAGCGCAGGCTGTACCGTGTTCCTTAGCTTGTGTGTGTGCGGCCCGT
GTGATGGGCTAGCCGGACCTGCTGGAGGCTGGTCCCTTTTGGGAAACCTGGAGCCAGGTGTACA
ACCCTTGCCATGAAGGGCCAGGATGCCAGATGCTTGGCCCCGTGAAGTGTGTCTGGAGCAG
CAGGATCCCGGGACAGGATGGGGGGCTTGGGGAAAACCTGCACTTCTGCACATTTGAAAAGAG
CAGCTGCTGCTTAGGGCCGCCGAAGCTGGTGTCTGTCAATTTCTCTCAGGAAAGGTTTTCAA
GTTCTGCCCATTTCTGGAGGCCACCACTCCTGTCTCTTCTCTTTCCCATCCCTGCTACCCCTG
GCCAGCACAGGCACTTTCTAGATATTTCCCCCTTGTGAGAGAAGAAAGAGCCCTGGTTTTATT
TGTTTGTACTACTCACTCAGTGAGCATCTACTTTGGGTGCATTCTAGTGTAGTTACTAGTCTT
TTGACATGGATGATTCTGAGGAGGAAGCTGTATTGAATGTATAGAGATTTATCCAATAAATAT
CTTTATTTAAAAATGAAAAA

FIGURE 156

MRERPRLGEDSSLISLFLQVVAFLAMVMGHTYSHWPSCCPSKGQDTSEELLRWSTVPVPPLEPA
RPNRHPESCRASEDGPLNSRAISPWRYELDRDLNRLPQDLYHARCLCPHCVSLQTGSHMDPRGNS
ELLYHNQTVFYRRPCHGEKGTGTHKGYCLERRLYRVSLACVCRPRVMG

Important features of the protein:

Signal peptide:

amino acids 1-32

N-glycosylation site.

amino acids 136-140

Tyrosine kinase phosphorylation site.

amino acids 127-135

N-myristoylation sites.

amino acids 44-50, 150-156

FIGURE 157

CCGGCGATGTCGCTCGTGCTGCTAAGCCTGGCCGCGCTGTGCAGGAGCGCCGTACCCCGAGAGCC
GACCGTTCAATGTGGCTCTGAAACTGGGCCATCTCCAGAGTGGATGCTACAACATGATCTAATCC
CCGGAGACTTGAGGGACCTCCGAGTAGAACCTGTTACAAC TAGTGTGCAACAGGGGACTATTCA
ATTTTGATGAATGTAAGCTGGGTACTCCGGGCAGATGCCAGCATCCGCTTGTGAAGGCCACCAA
GATTTGTGTGACGGGCAAAGCAACTTCCAGTCTACAGCTGTGTGAGGTGCAATTACACAGAGG
CCTTCCAGACTCAGACCAGACCCTCTGGTGGTAAATGGACATTTTCTTACATCGGCTTCCCTGTA
GAGCTGAACACAGTCTATTTCAATGGGGCCATAATATTCCTAATGCAAATATGAATGAAGATGG
CCCTTCCATGTCTGTGAATTTACCTCACCAGGCTGCCTAGACCACATAATGAAATATAAAAAAA
AGTGTGTCAAGGCCGGAAGCCTGTGGGATCCGAACATCACTGCTTGTGAAGAAGATGAGGAGACA
GTAGAAGTGAACCTTCAACACCCTCCCTGGGAAACAGATACATGGCTCTTATCCAACACAGCAC
TATCATCGGGTTTTCTCAGGTGTTTGAGCCACACCAGAAGAAACAAACGCGAGCTTCAGTGGTGA
TTCCAGTACTGAGGATAGTGAAGGTGCTACGGTGCAGCTGACTCCATATTTTCTACTTGTGGC
AGCGACTGCATCCGACATAAAGGAACAGTTGTGCTCTGCCACAAACAGGCGTCCCTTTCCCTCT
GGATAACAACAAAAGCAAGCCGGGAGGCTGGCTGCCTCTCCTCCTGCTGTCTGCTGGTGGCCA
CATGGGTGCTGGTGGCAGGGATCTATCTAATGTGGAGGCACGAAAGGATCAAGAAGACTTCCCTT
TCTACCACCACACTACTGCCCCCATTAAGGTTCTTGTGGTTTACCCATCTGAAATATGTTTCCA
TCACACAATTTGTTACTTCACTGAATTTCTTCAAACCATTGCAGAAGTGAGGTCATCCTTGAAA
AGTGGCAGAAAAAGAAAA TAGCAGAGATGGGTCCAGTGCAGTGGCTTGCCACTCAAAGAAGGCA
GCAGACAAAGTCGTCTTCTTCTTTCCAATGACGTCAACAGTGTGTGCGATGGTACCTGTGGCAA
GAGCGAGGGCAGTCCCAGTGAAGTCTCAAGACCTCTTCCCCCTTGCTTTAACCTTTTCTGCA
GTGATCTAAGAAGCCAGATTTCATCTGCACAAATACGTGGTGGTCTACTTTAGAGAGATTGATACA
AAAGACGATTACAATGCTCTCAGTGTCTGCCCCAAGTACCACCTCATGAAGGATGCCACTGCTTT
CTGTGCAGAACTTCTCCATGTCAAGCAGCAGGTGTCAGCAGGAAAAAGATCACAAGCCTGCCACG
ATGGCTGCTGCTCCTTGTAG

FIGURE 158

MSLVLLSLAALCRSAVPREPTVQCGSETGPSPEWMLQHDLPDGLRDLRVEPVTTSVATGDYSILMNVS
WRADASIRLLKATKICVTGKSNFQSYSCVRCNYTEAFQTQTRPSGGKWTFSYIGFPELNTVYFIGAHNIP
NANMNEDGPSMSVNFTSPGCLDHIMKYKKKCVKAGSLWDPNITACKNEETVEVNFTTTPLGNYMALIQH
STIIGFSQVFEHQKQTRASVVIPVTGDSEGATVQLTPYFPTCGSDCIRHKGTVVLCPTGVPFPLDNNK
SKPGGWLPLLLLSLLVATWVLVAGIYLMWRHERIKKTSFSTTTLLPPIKVLVVYPSEICFHHTICYFTEFL
QNHCRSEVILEKWQKKKIAEMGPVQWLATQKKAADKVVFLSNDVNSVCDGTCGKSEGSSENSQDLFPLA
FNLFCSDLRSQIHLHKYVVVYFREIDTKDDYNALSVC PKYHLMKDATAFCAELLHVKQVVSAGKRSQACHD
GCCSL

Important features of the protein:

Signal peptide:

amino acids 1-14

Transmembrane domain:

amino acids 290-309

N-glycosylation sites.

amino acids 67 - 71, 103 - 107, 156 - 160, 183 - 187, 197 - 201 and 283
- 287

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 228 - 232 and 319 - 323

Casein kinase II phosphorylation sites.

amino acids 178 - 182, 402 - 406, 414 - 418 and 453 - 457

N-myristoylation site.

amino acids 116-122

Amidation site.

amino acids 488-452

FIGURE 159

AGCCACCAGCGCAACATGACAGTGAAGACCCTGCATGGCCAGCCATGGTCAAGTACTTGCTGCT
GTCGATATTGGGGCTTGCCCTTTCTGAGTGAGGCGGCAGCTCGGAAAATCCCCAAAGTAGGACATA
CTTTTTTCCAAAAGCCTGAGAGTTGCCCGCCTGTGCCAGGAGGTAGTATGAAGCTTGACATTGGC
ATCATCAATGAAAACCAGCGCGTTTCCATGTCACGTAACATCGAGAGCCGCTCCACCTCCCCCTG
GAATTACACTGTCAC TTGGGACCCCAACCGGTACCCTCGGAAGTTGTACAGGCCCAGTGTAGGA
ACTTGGGCTGCATCAATGCTCAAGGAAAGGAAGACATCTCCATGAATTCCGTCCCATCCAGCAA
GAGACCCTGGTCGTCCGGAGGAAGCACCAAGGCTGCTCTGTTTCTTTCCAGTTGGAGAAGGTGCT
GGTGACTGTTGGCTGCACCTGCGTCAACCCTGTCATCCACCATGTGCAGTAAGAGGTGCATATCC
ACTCAGCTGAAGAAG

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

FIGURE 160

MTVKTLHGPMVKYLLLSILGLAFLSEAAARKIPKVGHTFFQKPESCPPVPGGSMKLDIGIINEN
QRVSMRNIESRSTSPWNYTVTWDPNRYPSEVVQAQCRNLGCINAQGKEDISMNSVPIQQETLVV
RRKHQGCVSFQLEKVLVTVGCTCVTPVIHHVQ

Signal sequence:

amino acids 1-30

N-glycosylation site.

amino acids 83-87

N-myristoylation sites.

amino acids 106-111, 136-141

FIGURE 161

ACACTGGCCAAACAAAAACGAAAGCACTCCGTGCTGGAAGTAGGAGGAGAGTCAAGACTCCCAGG
ACAGAGAGTGCACAACTACCCAGCACAGCCCCCTCCGCCCTTGGAGGCTGAAGAGGGATTC
CAGCCCCGCCACCCACAGACACGGGTGACTGGGTGTCTGCCCTTGGGGGGGGCAGCAC
AGGGCCTCAGGCCTGGGTGCCACCTGGCACCTAGAAAGATGCCTGTGCCCTGGTCTTGCTGTCTT
TGGCACTGGGCCGAAGCCAGTGGTCTTTCTCTGGAGAGGCTTGTGGGGCCTCAGGACGCTACC
CACTGCTCTCCGGGCTCTCCTGCCGCTCTGGGACAGTACATACTCTGCCTGCCTGGGGACAT
CGTGCCTGCTCCGGGCCCCGTGCTGGGCTACGCACCTGCAGACAGAGCTGGTGTGAGGTGCC
AGAAGGAGACCGACTGTGACCTCTGTCTGCGTGTGGCTGTCCACTTGGCCGTGCATGGGCACTGG
GAAGAGCCTGAAGATGAGGAAAAGTTTGGAGGAGCAGTCACTCAGGGGTGGAGGAGCCTAGGAA
TGCTCTCTCCAGGCCAAGTCTGCTCTCCTTCCAGGCCTACCTACTGCCCGCTGCGTCTGCTG
TGGAGGTGCAAGTGCCTGCTGCCCTTGTGCAATTTGGTCACTGTGGGCTCTGTTGATATGAC
TGCTTCGAGGCTGCCCTAGGGAGTGGGTACGAATCTGGTCTATACTCAGCCCAGGTACGAGAA
GGAATCAACACACACAGCAGTGCCTGCCCTGCCCTGGCTCAACGTGTCAAGCAGATGGTGA
ACGTGCATCTGGTTCTGAATGTCTCTGAGGAGCAGCACTTCGGCCTCTCCCTGTACTGGAATCAG
GTCCAGGGCCCCCAAAACCCCGGTGGCACAAAACCTGACTGGACCGCAGATCATTACCTTGAA
CCACACAGACCTGGTTCCTGCTCTGTATTAGGTGTGGCCTCTGGAACCTGACTCCGTTAGGA
CGAACATCTGCCCTTCCAGGGAGACCCCGCGCACACCAGAACCTCTGGCAAGCCGCCGACTG
CGACTGCTGACCTGCAGAGTGGCTGCTGGACGCACCGTGTGCTGCTGCCCGCAGAAGCGGCACT
GTGCTGGCGGGCTCCGGGTGGGGACCCCTGCCAGCCACTGGTCCCACCGCTTCTCTGGGAGAACG
TCACTGTGGACAAGTTCTCGAGTTCCTTGTGAAAGGCCACCCTAACCTCTGTGTTCAAGTG
AACAGCTCGGAGAAGCTGCAGCTGCAGAGTGTCTTGTGGGCTGACTCCCTGGGGCCTCTCAAAGA
CGATGTGCTACTGTTGGAGACACGAGGCCCCAGGACAACAGATCCCTCTGTGCTTGGAAACCCA
GTGGCTGTACTTCACTACCCAGCAAAGCCTCCACGAGGGCAGCTCGCCTTGGAGAGTACTTACTA
CAAGACCTGCAGTCAAGCCAGTGTCTGCAGCTATGGGACGATGACTTGGGAGCGCTATGGGCTG
CCCCATGGACAAATACATCCACAAGCGCTGGGCCCTCGTGTGGCTGGCCTGCCTACTCTTTGCCG
CTGGCCTTCCCTCATCTCCTTCTCAAAAAGGATCACGCGAAAGGGTGGCTGAGGCTCTTGAAA
CAGGACGTCCGCTCGGGGGCGGCCGAGGGGGCGCGGCTCTGCTCCTACTCAGCCGATGA
CTCGGGTTTCGAGCGCCTGGTGGGCGCCTGGCGTCCGCCCTGTGCCAGCTGCCGCTGCCGCTGG
CCGTAGACCTGTGGAGCCGTGCTGAAGTGAAGCGCAGGGGCCCGTGGCTTGGTTTACGCGCAG
CGGCGCCAGACCCTGCAGGAGGGCGGGGTGGTGGTCTTGTCTTCTCTCCCGGTGCGGTGGCGCT
GTGCAGCGAGTGGCTACAGGATGGGTGTCCGGGCCGGGGCGCACGCCCCGACGACGCTTCC
GCGCCTCGCTCAGTGCCTGCTGCCGACTTCTTGCAGGGCCGGGCGCCCGGACGCTACGTGGGG
GCCTGCTTCGACAGGTGCTCCACCCGGACGCCGTACCCGCCCTTTCCGCACCGTGCCTGCTT
CACACTGCCCTCCCAACTGCCAGACTTCTTGGGGCCCTGCAGCAGCCTCGCGCCCCGCTTCCG
GGCGGCTCCAAGAGAGAGCGGAGCAAGTGTCCCGGGCCCTTACGCCAGCCCTGGATAGCTACTTC
CATCCCCGGGGACTCCCGCGCCGGGACGCGGGTGGGACCAGGGGGCGGACCTGGGGCGGGGA
CGGGACTTAAATAAAGGCAGACGCTGTTTTCTAAAAAA

FIGURE 162

MPVPWFLLSLALGRSPVVLSELERLVGPQDATHCSPGLSCRLWDSILCLPGDIVPAPGPVLAPTHLQTELV
LRCQKETDCDLCLRVAVHLAVHGHWEPEDEEKFGGAADSGVEEPRNASLQAQVVLSFQAYPTARCVLLEV
QVPAALVQFGQSVGSVVYDCFEAALGSEVRIWSYTPRYEKELNHTQQLPALPWLNVVSADGDNVHLVNLVS
EEQHFGLSLYWNQVQGGPKRWHKNLTGPQIITLNHTDLVPCLCIQWVPLEPDSVRTNICPFREDPRAHQN
LWQAARLRLLLTQSWLLDAPCSLPAAALCWRAPGGDPCQPLVPPLSWENVTVDKVLEFLLKGHPNLCVQ
VNSSEKLQEQECLWADSLGPLKDDVLLLETRGPDNRSLEPSGCTSLPSKASTRAARLGEYLLQDLQS
GQCLQLWDDDLGALWACPMCKYIHKRWALVWLACLLFAAALSLLLLKDHAKGWLRLKQDVRSGAAARG
RAALLYSADDSGFERLVGALASALCQLPLRVAVDLWSRRELSAQGPVAVFHAQRRQTLQEGGVVLLFSP
GAVALCSEWLQDGVSGPAGHGHDAFRASLSCVLPDFLQGRAPGSYVACFDRLHHPDAVPALEFRTVPVFT
LPSQLPDFLQALQPRAPRSRGLQERAEQVSRALQPALDSYFHPPGTPAPGRGVGPGAGPGAGDGT

Signal sequence:

amino acids 1-20

Transmembrane domain.

amino acids 453-475

N-glycosylation sites.

amino acids 118-121, 186-189, 198-201, 211-214, 238-241, 248-251,
334-337, 357-360, 391-394

Glycosaminoglycan attachment site.

amino acids 583-586

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 552-555

N-myristoylation sites.

amino acids 107-112, 152-157, 319-324, 438-443, 516-521, 612-617,
692-697, 696-701, 700-705

FIGURE 163

GGGAGGGCTCTGTGCCAGCCCCGATGAGGACGCTGCTGACCATCTTGACTGTGGGATCCCTGGCT
GCTCAGCCCCCTGAGGACCCCTCGGATCTGCTCCAGCACGTGAAATCCAGTCCAGCAACTTTGA
AAACATCCTGACGTGGGACAGCGGGCCAGAGGGCACCCAGACACGGTCTACAGCATCGAGTATA
AGACGTACGGAGAGAGGGACTGGGTGGCAAAGAAGGGCTGTCAGCGGATCACCCGGAAGTCTGC
AACCTGACGGTGGAGACGGGCAACCTCACGGAGCTCTACTATGCCAGGGTCAACCGT
GTCAGTGGGGAGGGCCGGTACGCCACCAAGATGACTGACAGGTTGAGCTCTCTGCAGCACACTAC
CCTCAAGCCACCTGATGTGACCTGTATCTCCAAAGTGAGATCGATTGAGATGATTGTTTCATCTA
CCCCACGCCAATCCGTGCAGGCGATGGCCACCGCTAACCTGGAGACATCTCCATGACCTG
TTCTACCACCTAGAGCTCCAGGTCACCCGACCTACCAAATGCACCTGGAGGGGAGCAGAGAGA
ATATGAGTTCTTCGGCTGACCCCTGACACAGAGTTCCTGGCACCATCATGATTTGCGTTCCCA
CCTGGGCCAAGGAGAGTGGCCCTACATGTGCCGAGTGAAGACACTGCCAGACCCGACATGGACC
TACTCCTTCCCGGAGCCTTCTGTTCTCCATGGGCTTCTCGTCCGAGTACTCTGCTACCTGAG
CTACAGATATGTCACCAAGCCGCTGCACCTCCAACTCCCTGAACGTCCAGCGAGTCTGACTT
TCCAGCCGCTGCGCTTCATCCAGGACGTCCTGATCCCTGTCTTTGACCTCAGCGGCCACGAGC
AGTCTGGCCAGCCTGTCCAGTACTCCAGATCAGGGTGTCTGGACCCAGGGAGCCCGCAGGAGC
TCCACAGCGGCATAGCCTGTCCGAGATCACCTACTTAGGGCAGCCAGACATCTCCATCCTCCAGC
CCTCCAACGTGCCACCTCCCGAGATCCTTCCCCACTGTCTATGCCCAAACGCTGCCCTGAG
GTCGGGCCCCCATCCTATGCACCTCAGGTGACCCCGAAGCTCAATGCCATTCTACGCCCCACA
GGCCATCTTAAGGTCCAGCCTTCTCCTATGCCCTCAAGCCACTCCGGACAGCTGGCCTCCCT
CCTATGGGGTATGCAATGGAAGTTCTGGCAAAGACTCCCCACTGGGACACTTTCTAGTCTTAA
CACCTTAGGCCTAAAGGTGAGCTTTCAGAAAGAGCCACCAGTGGAGCTGCATGTTAGGTGGCT
TTCTCTGCAGGAGGTGACCTCCTTGGCTATGGAGGAATCCCAAGAAGCAAAATCATTGCACCAGC
CCCTGGGGATTTCACAGACAGAACATCTGACCCAAATGTGCTACACAGTGGGGAGGAAGGGACA
CCACAGTACCTAAAGGGCAGCTCCCCCTCCTCTCCTCAGTCCAGATCGAGGGCCACCCCATGTC
CCTCCCTTTGCAACCTCCTTCCGGTCCATGTTCCCCCTCGGACCAAGGTCCAAGTCCCTGGGGCC
TGCTGGAGTCCCTTGTGTGTCCCAAGGATGAAGCCAAGAGCCAGCCCTGAGACCTCAGACCTG
GAGCAGCCACAGAATGGATTCTCTTTTCAGAGGCTGGCCCTGACTGTGCAGTGGGAGTCTTG
AGGGGAATGGGAAAGGCTTGGTGCTTCTCCCTGTCCCTACCCAGTGTACATCCTTGGCTGTCA
ATCCCATGCCTGCCCATGCCACACACTCTGCGATCTGGCCTCAGACGGGTGCCCTTGAGAGAAGC
AGAGGGAGTGGCATGCAGGGCCCTGCCATGGGTGCGCTCCTCACCGGAACAAGCAGCATGATA
AGGACTGCAGCGGGGGAGCTCTGGGGAGCAGTTGTGTAGACAAGCGCGTGTCTGCTGAGCCCTG
CAAGGCAGAAATGACAGTGAAGGAGGAAATGCAGGGAAACTCCCGAGGTCCAGAGCCCCACCTC
CTAACACCATGGATTCAAAGTGCTCAGGGAATTTGCCTCTCCTTGCCCCATTCTTGCCAGTTTC
ACAATCTAGCTCGACAGAGCATGAGGCCCTGCCCTTCTGTCTATTGTTCAAAGGTGGGAAGAGA
GCCTGGAAAAGAACCAGGCTGGAAAAGAACCAGAAGGAGGCTGGGCAGAACCAGAACAACCTGC
ACTTCTGCCAAGGCCAGGGCCAGCAGGACGGCAGGACTCTAGGGAGGGGTGTGGCTGCAGCTCA
TTCCAGCCAGGGCAACTGCCTGACGTTGCACGATTTAGCTTCATTCCTCTGATAGAACAAGC
GAAATGCAGGTCCACCAGGGAGGGAGACACACAAGCCTTTTCTGCAGGCAGGAGTTTACAGCCCT
ATCCTGAGAATGGGGTTTGAAGGAAGGTGAGGGCTGTGGCCCTGGACGGGTACAATAACACAC
TGTAAGTGTACAACCTTGAAGCTCTGCCCTGGGTTGAGCCATCTGGGCTCAAATTCAGC
CTCACCCTCACAAGCTGTGTGACTTCAAACAAATGAAATCAGTGCCAGAACCTCGGTTTCTC
ATCTGTAATGTGGGATCATAACACCTACCTCATGGAGTTGTGGTGAAGATGAAATGAAGTCATG
TCTTTAAAGTGCTTAATAGTGCCTGGTACATGGGCAGTGCCCAATAAACGGTAGCTATTTAAAAA
AAAAAAA

10063728

FIGURE 164

MRTLLTILTVGSLAAHAPEDPSDLLQHVKFQSSNFENILTWDSGPEGTPDTVYSIEYKTYGERDW
VAKKGCQRITRKSCNLTVE TGNLTEL Y YARVTAVSAGGRSATKMTDRFSSLQHTTLKPPDVTCIS
KVRSIQMI VHP TPTPIRAGDGHRLTLEDI FHDLFYHLELQVNRTYQMHLGGKQREYEFFGLTPDT
EFLGTIMICVPTWAKESAPYMCRVKTL PDRTWTYSFSGAFLFSMGFLVAVLCYLSYRYVTKPPAP
PNSLNVQRVLT FQPLRFIQEHVLI PVFDLSGPSSLAQPVQYSQIRVSGPREPAGAPQRHSLSEIT
YLGQPDISILQPSNVPPPQILSPLSYAPNAAPEVGPSSYAPQVTPEAQFPFYAPQAISKVQPSSY
APQATPDSWPPSYGVCMEGSGKDSPTGTLSSPKHLRPGQLQKEPPAGSCMLGGLSLQEVTSLAM
EESQEA KSLHQPLGICTDRTSDPNV LHSGEETPQYLKQLPLLSSVQIEGH PMSLPLQPPSGPC
SPSDQGSPWGLLES LVC PKDEAKSPAPETS DLEQPT ELDSLFRGLALT VQWES

Signal sequence.

amino acids 1-17

Transmembrane domain.

amino acids 233-250

N-glycosylation sites.

amino acids 80-83, 87-90, 172-175

N-myristoylation sites.

amino acids 11-16, 47-52, 102-107, 531-536, 565-570

FIGURE 165

TGGCCTACTGGAAAAAAAAAAAAAAAAAAAAAGTCACCCGGGCCCGCGGTGGCCACAACATGG
CTGCGGGCGCCGGGGCTGCTCTTCTGGCTGTTTCGTGCTGGGGGGCGCTCTGGTGGGTCCCAGGGCCAG
TCGGATCTCAGCCACGGACGGCGTTTTCTCGGACCTCAAAGTGTGCGGGGACGAAGAGTGCAGCAT
GTTAATGTACCGTGGGAAAGCTCTTGAAGACTTCACGGGCCCTGATTGTCGTTTTGTGAATTTA
AAAAAGGTGACGATGTATATGTCTACTACAACTGGCAGGGGGATCCCTTGAACCTTGGGCTGGA
AGTGTGGAACACAGTTTTGGATATTTCCAAAAGATTTGATCAAGGTACTTCATAAATACACGGA
AGAAGAGCTACATATCCAGCAGATGAGACAGACTTTGTCTGCTTTGAAGGAGGAAGAGATGATT
TTAATAGTTATAATGTAGAAGAGCTTTAGGATCTTTGGAAGTGGAGACTCTGTACCTGAAGAG
TCGAAGAAAGCTGAAGAAGTTTCTCAGCACAGAGAGAAATCTCCTGAGGAGTCTCGGGGGCGTGA
ACTTGACCTGTGCCTGAGCCCAGGCATTCAGAGCTGATTCAGAGGATGGAGAAGGTGCTTTCT
CAGAGAGCACCGAGGGCTGCAGGGACAGCCCTCAGCTCAGGAGAGCCACCCTCACACCAGCGGT
CCTGCGGCTAACGCTCAGGGAGTGCAGTCTTCGTTGGACACTTTTGAAGAAATCTGCACGATAA
ATTGAAAGTGCCGGGAAGCGAAAGCAGAACTGGCAATAGTTCTCCTGCCTCGGTGGAGCGGGAGA
AGACAGATGCTTACAAAGTCTGAAAACAGAAATGAGTCAGAGAGGAAGTGGACAGTGCCTTATT
CATTACAGCAAAGATTCGTTGGCATCAAATCTAAGTTTGTTTTACAAAGATTGTTTTTAGTA
CTAAGCTGCCTTGGCAGTTGCATTTTTGAGCCAAACAAAAATATATTATTTCCCTTCTAAGTA
AAAAAAAAAAAAAAAAAAAA

FIGURE 166

MAAAPGLLFWLFLVGLALWWVPGQSDLSHGRRFSDLKVCGDEECMSMLMYRGKALEDFTPDCRFVN
FKKGDDVYVYKLAGGSLELWAGSVEHSFGYFPKDLIKVLHKYTEELHIPADETDFVCFEGGRD
DFNSYNVEELGSLLELDSVP EESKKAEEV SQHREKSP EESRGRELDPVPEPEAFRA DSEGE GA
FSESTEGLQGQPSAQESH PHTSGPAANAQGVQSSLDTFEEILHDKLKVPGSESRTGNSSPASVER
EKTDAYKVLKTEMSQRGSGQCVIHYSKGFRWHQNL SLFYKDCF

Important features of the protein:

Signal peptide:

amino acids 1-22

N-glycosylation site.

amino acids 294-298

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 30-34

Tyrosine kinase phosphorylation site.

amino acids 67-76

N-myristoylation sites.

amino acids 205-211, 225-231, 277-283

Amidation site.

amino acids 28-32

FIGURE 167

CCAGGACCAGGGCGCACCGGCTCAGCCTCTCACTTGTTCAGAGGCCGGGGAAGAGAAGCAAAGCGC
AACGGTGTGGTCCAAGCCGGGGCTTCTGCTTCGCCTCTAGGACATACACGGGACCCCTAACTTC
AGTCCCCAAACGCGCACCCCTCGAAGTCTTGAAGTCCAGCCCCGCACATCCACGCGCGGCACAGG
CGCGGCAGGCGGCAGGTCCCGGCCGAAGGCGATGCGCGCAGGGGGTTCGGGCAGCTGGGCTCGGGC
GGCGGGAGTAGGGCCCCGGCAGGGAGGCAGGGAGGCTGCATATTCAGAGTCGCGGGCTGCGCCCTG
GGCAGAGGCCGCCCTCGCTCCACGCAACACCTGCTGCTGCCACCGCGCCCGGATGAGCCGCGTGG
TCTCGCTGCTGCTGGGCGCCGCGCTGCTCTGCGGCCACGGAGCCTTCTGCCGCCGCGTGGTTCAGC
GGCCAAAAGGTGTGTTTTGCTGACTTCAAGCATCCCTGCTACAAAATGGCCTACTTCCATGAACT
GTCCAGCCGAGTGAGCTTTTACAGAGGCACGCCTGGCTTGTGAGAGTGAGGGAGGAGTCTCCTCA
GCCTTGAGAATGAAGCAGAACAGAAGTTAATAGAGAGCATGTTGCAAAACCTGACAAAACCCGGG
ACAGGGATTTCTGATGGTGATTTCTGGATAGGGCTTTGGAGGAATGGAGATGGGCAAACATCTGG
TGCCTGCCCAGATCTCTACCAGTGGTCTGATGGAAGCAATTCCCAGTACCGAAACTGGTACACAG
ATGAACCTTCCCTGCGGAAGTGAAAAGTGTGTTGTGATGTATCACCAACCAACTGCCAATCCTGGC
CTTGGGGTCCCTACCTTTTACCAGTGGAAATGATGACAGGTGTAACATGAAGCACAATTATATTTG
CAAGTATGAACCAGAGATTAATCCAACAGCCCCTGTAGAAAAGCCTTATCTTACAAATCAACCAG
GAGACCCCATCAGAATGTGGTGTACTGAAGCAGGTATAATTCCCAATCTAATTTATGTTGTT
ATACCAACAATACCCCTGCTCTTACTGATACTGGTTGCTTTTTGGAACCTGTTGTTTTCCAGATGCT
GCATAAAAGTAAAGGAAGAACAATAACTAGTCCAAACCAGTCTACACTGTGGATTTCAAAGAGTA
CCAGAAAAGAAAGTGGCATGGAAGTATAATAACTCATTGACTTGGTTCAGAAATTTTGTAAATCT
GGATCTGTATAAGGAATGGCATCAGAACAATAGCTTGAATGGCTTGAAATCACAAAGGATCTGC
AAGATGAACTGTAAGTCCCCCTTGGAGCAAATATTAAGTAATTTTTATATGTCTATTTATTTCA
TTTAAAGAATATGCTGTGCTAATAATGGAGTGAGACATGCTTATTTTGCTAAAGGATGCACCCAA
ACTTCAAACCTCAAGCAAATGAAATGGACAATGCAGATAAAGTTGTTATCAACACGTCGGGAGTA
TGTGTGTAGAAAGCAATTCCTTTTATTTCTTTACCTTTTATAAGTTGTTATCTAGTCAATGTAA
TGTATATTTGATTGAAATTTACAGTGTGCAAAAGTATTTTACCTTTGCATAAGTGTGATAAAA
ATGAACTGTTCTAATATTTATTTTATGGCATCTCATTTTTCAATACATGCTCTTTTGATTAAAG
AACTTATTACTGTTGTCAACTGAATTCACACACACACAAATATAGTACCATAGAAAAAGTTTGT
TTTCTCGAAATTAATCATCTTTCAGCTTCTCTGCTTTTGGTCAATGTCTAGGAAATCTCTTCAGA
ATAAGAAGCTATTTCAATTAAGTGTGATATAAACCTCCTCAAACATTTTACTTAGAGGCAAGGAT
TGTCTAATTTCAATTTGCAAGACATGTGCCTTATAATTTTTTAGCTTAAAATTAACAGATT
TTGTAATAATGTAACCTTTGTTAATAGGTGCATAAACACTAATGCAGTCAATTTGAACAAAAGAAG
TGACATACACAATATAAATCATATGTCTTACACGTTGCCATATAATGAGAAGCAGCTCTCTGA
GGTTCTGAAATCAATGTGGTCCCTCTCTTGGCCACTAAACAAAGATGGTTGTTTCGGGGTTTGGG
ATTGACACTGGAGGCAGATAGTTGCAAAGTTAGTCTAAGGTTTCCCTAGCTGTATTTAGCCTCTG
ACTATATTAGTATACAAAGAGGTCATGTGGTTGAGACCAGGTGAATAGTCACTATCAGTGTGGAG
ACAAGCACAGCACACAGACATTTTAGGAAGGAAAGGAACACGAAATCGTGTGAAAATGGGTTGG
AACCCATCAGTGATCGCATATTCATTGATGAGGGTTTGCTTGGATAGAAAATGGTGGCTCCTTT
CTGTCTTATCTCCTAGTTTCTTCAATGCTTACGCCTTGTCTTCTCAAGAGAAAGTTGTAACCTCT
CTGGTCTTCATATGTCCCTGTGCTCCTTTTAAACCAATAAAGAGTCTTGTCTTCTGGGGGAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

FIGURE 168

MSRVVSLLLGAALLCGHGAFRRVSVSGQKVCFADEFKHPCYKMAYFHELSSRVSFQEARLACESE
GGVLLSLENEAEQKLIESMLQNLTKPGTGISDGDWIGLWRNGDGQTSACPDLYQWSDGSNSQ
YRNWYTDEPSCGSEKCVVMYHQPTANPGLGGPYLYQWDDRCNMKHNYICKYEPEINPTAPVEK
PYLTNQPGDTHQNVVVTEAGIIPNLIYVVIPTIPLLLLILVAFGTCCFQMLHKSkgRtKtSPNQ
STLWISKSTRKESGMEV

Important features of the protein:

Signal peptide:

amino acids 1-21

Transmembrane domain:

amino acids 214-235

N-glycosylation sites.

amino acids 86-89, 255-258

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 266-269

N-myristoylation sites.

amino acids 27-32, 66-71, 91-96, 93-98, 102-107, 109-114, 140-
145, 212-217